

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 15 Seconds

(without alignments)  
262.845 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKVTVAENQGPNNRVFIKR.....KFIKRVGIIAAYQNPASWK 134

Scoring table: BLOSUM62

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/aa/5A COMB.pep:\*

2: /cgn2\_6/ptodata/1/aa/5B COMB.pep:\*

3: /cgn2\_6/ptodata/1/aa/6A COMB.pep:\*

4: /cgn2\_6/ptodata/1/aa/6B COMB.pep:\*

5: /cgn2\_6/ptodata/1/aa/PCITUS COMB.pep:\*

6: /cgn2\_6/ptodata/1/aa/backfillseq.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	26.4	60	3	US-08-467-023-62
2	188	26.4	374	3	US-08-467-023-2
3	150	21.1	367	3	US-08-467-023-95
4	150	21.1	370	3	US-08-467-023-97
5	148.5	20.9	514	3	US-08-467-023-134
6	115	16.2	128	3	US-08-467-023-187
7	105	14.8	20	3	US-08-467-023-33
8	99	13.9	127	3	US-08-467-023-188
9	98.5	13.9	387	1	US-08-290-448A-72
10	98.5	13.9	387	1	US-08-290-448A-72
11	98.5	13.9	387	1	US-08-175-069A-72
12	98.5	13.9	387	1	US-08-461-939B-72
13	98.5	13.9	387	1	US-08-461-939B-76
14	98.5	13.9	387	1	US-08-290-448A-76
15	97	13.6	397	1	US-08-290-448A-76
16	97	13.6	397	1	US-08-175-069A-76
17	97	13.6	397	1	US-08-461-939B-76
18	97	13.6	397	1	US-08-461-939B-76
19	93	13.1	391	1	US-08-290-448A-59
20	93	13.1	391	1	US-08-290-448A-59
21	93	13.1	391	1	US-08-175-069A-59
22	93	13.1	391	1	US-08-461-939B-59
23	93	13.1	391	1	US-08-461-939B-59
24	93	13.1	398	1	US-08-290-448A-74
25	93	13.1	398	1	US-08-290-448A-74
26	93	13.1	398	1	US-08-175-069A-74
27	93	13.1	398	1	US-08-461-939B-74

28	93	13.1	398	4	US-08-464-000-74	Sequence 74, Appl
29	91.5	12.9	30	3	US-08-467-023-71	Sequence 71, Appl
30	91	12.8	20	3	US-08-467-023-34	Sequence 34, Appl
31	88	12.4	383	1	US-08-290-448A-78	Sequence 78, Appl
32	88	12.4	383	1	US-08-290-448A-78	Sequence 78, Appl
33	88	12.4	383	1	US-08-175-069A-78	Sequence 78, Appl
34	88	12.4	383	1	US-08-461-939B-78	Sequence 78, Appl
35	88	12.4	383	4	US-08-464-000-80	Sequence 80, Appl
36	86	12.1	388	1	US-08-290-448A-80	Sequence 80, Appl
37	86	12.1	388	1	US-08-175-069A-80	Sequence 80, Appl
38	86	12.1	388	1	US-08-461-939B-80	Sequence 80, Appl
39	86	12.1	388	4	US-08-464-000-80	Sequence 80, Appl
40	83.5	11.7	30	3	US-08-467-023-67	Sequence 67, Appl
41	81.5	11.5	24	4	US-08-467-023-72	Sequence 72, Appl
42	81.5	11.5	234	4	US-09-219-983A-7	Sequence 7, Appl
43	81.5	11.5	20	3	US-08-467-023-32	Sequence 32, Appl
44	81	11.4	127	3	US-08-467-023-189	Sequence 189, Appl
45	81	11.4				

## ALIGNMENTS

## RESULT 1

US-08-467-023-62  
Sequence 62, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garmen, Richard D.;

APPLICANT: Kuo, Wei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.

TITLE OF INVENTION: Allergenic Proteins And Peptides From

TITLE OF INVENTION: Japanese Cedar Pollen

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (INT-028CPD2)

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 62:

SEQUENCE CHARACTERISTICS:

LENGTH: 60 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-467-023-62

Query Match 26.4%; Score 188; DB 3; Length 60;  
Best Local Similarity 71.2%; Pred. No. 1.5e-17;  
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 73 RPLWIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 124  
Db 6 RPLWIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 57

RESULT 2

US-08-467-023-2

Sequence 2, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garman, Richard D.;

APPLICANT: Kuo, Mei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.;

TITLE OF INVENTION: Allergenic Proteins And Peptides From

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESS: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 374 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-467-023-2

Query Match 26.4%; Score 188; DB 3; Length 374;  
Best Local Similarity 71.2%; Pred. No. 1.9e-16;  
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 73 RPLWIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 124  
Db 87 RPLWIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 138

RESULT 3

US-08-467-023-95

Sequence 95, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garman, Richard D.;

APPLICANT: Kuo, Mei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.;

TITLE OF INVENTION: Allergenic Proteins And Peptides From

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESS: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,023

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 367 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-467-023-95

Query Match 21.1%; Score 150; DB 3; Length 367;  
Best Local Similarity 55.8%; Pred. No. 1.9e-11;  
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLWIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 124  
Db 87 KALWIFSGNNIKLKMPLYAGYKTFDGRRAEVSYYHVNAGKFIIRVDGII 138

RESULT 4

US-08-467-023-97

Sequence 97, Application US/08467023

Patent No. 6090386

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.;

APPLICANT: Garman, Richard D.;

APPLICANT: Kuo, Mei-Chang;

APPLICANT: Yeung, Siu-mei H.;

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.;

APPLICANT: Powers, Steven P.;

TITLE OF INVENTION: Allergenic Proteins And Peptides From

US-08-467-023-95

TITLE OF INVENTION: Japanese Cedar Pollen  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,023  
FILING DATE: June 6, 1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane E. Remillard  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMT-028CPD2)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 97:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 370 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-467-023-97

Query Match 21.1%; Score 150; DB 3; Length 370;  
Best Local Similarity 55.8%; Pred. No. 1.9e-11;  
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMTISGNNNITLKMPTIAGYKTPDGRRAEVSYYHVNGAKFIRRVDCII 124  
DB 87 KALWIFSONNNIKLKMPLVYAGHKTIDRGADVHLNGGRCPLFMRYVSHVI 138

RESULT 5  
US-08-467-023-134  
Sequence 134, Application US/08467023  
Patent No. 6090386

GENERAL INFORMATION:  
APPLICANT: Griffith, Irwin J.;  
APPLICANT: Pollock, Joanne;  
APPLICANT: Bond, Julian F.;  
APPLICANT: Garman, Richard D;  
APPLICANT: Kuo, Wei-Chang;  
APPLICANT: Yeung, Siu-mei H.;  
APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powers, Steven P.  
TITLE OF INVENTION: Allergenic Proteins And Peptides From  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:

SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,023  
FILING DATE: June 6, 1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane E. Remillard  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMT-028CPD2)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 134:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 514 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-467-023-134

Query Match 20.9%; Score 148.5; DB 3; Length 514;  
Best Local Similarity 41.4%; Pred. No. 4.8e-11;  
Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTRKNNRIMLOFAKLTGFLMGRRPLMTIFSGNNITLKM 90  
DB 236 IDIFASKNFHLQKNTIGTRKNNRIMLOFAKLTGFLMGRRPLMTIFSGNNITLKM 90

QY 91 KYIAGYKTPD-----RAEVSYYHVNGAKFIRRVDCII 123  
DB 273 ICGPHGISISLGRNRAEVSYYHVNGAKFIRRVDCII 311

RESULT 6  
US-08-467-023-187  
Sequence 187, Application US/08467023  
Patent No. 6090386

GENERAL INFORMATION:  
APPLICANT: Griffith, Irwin J.;  
APPLICANT: Pollock, Joanne;  
APPLICANT: Bond, Julian F.;  
APPLICANT: Garman, Richard D;  
APPLICANT: Kuo, Wei-Chang;  
APPLICANT: Yeung, Siu-mei H.;  
APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powers, Steven P.  
TITLE OF INVENTION: Allergenic Proteins And Peptides From  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:

350 ID NO. 131

NAME: Jane E. Remillard  
 REGISTRATION NUMBER: 38,872  
 REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941  
 INFORMATION FOR SEQ ID NO: 187:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 128 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: internal  
 US-08-467-023-187

Query Match 16.2%; Score 115; DB 3; Length 128;  
 Best Local Similarity 95.2%; Pred. No. 1.8e-07;  
 Matches 20; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 WKNNRIWQFAKLGFTLMGR 72  
 DB 98 WKNNRIWQFAKLGFTLMGK 118

# RESULT 7

US-08-467-023-33  
 Sequence 33, Application US/08467023  
 Patent No. 6090386

GENERAL INFORMATION:  
 APPLICANT: Griffith, Irwin J.;  
 APPLICANT: Pollock, Joanne;  
 APPLICANT: Bond, Julian F.;  
 APPLICANT: Garman, Richard D;  
 APPLICANT: Kuo, Mei-Chang;  
 APPLICANT: Yeung, Siu-mei H.;  
 APPLICANT: Brauer, Andrew;  
 APPLICANT: Exley, Mark A.;  
 APPLICANT: Powers, Steven P.  
 TITLE OF INVENTION: Allergenic Proteins And Peptides From  
 TITLE OF INVENTION: Japanese Cedar Pollen  
 NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
 STREET: 610 Lincoln St  
 CITY: Waltham  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02154

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/467,023  
 FILING DATE: June 6, 1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/350,225  
 FILING DATE: December 6, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872  
 REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 33:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 20 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

FRAGMENT TYPE: internal  
 US-08-467-023-33

Query Match 14.8%; Score 105; DB 3; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 IFSGNNNITKLKMPYIAGYK 97  
 DB 1 IFSGNNNITKLKMPYIAGYK 20

# RESULT 8

US-08-467-023-188  
 Sequence 188, Application US/08467023  
 Patent No. 6090386

GENERAL INFORMATION:  
 APPLICANT: Griffith, Irwin J.;  
 APPLICANT: Pollock, Joanne;  
 APPLICANT: Bond, Julian F.;  
 APPLICANT: Garman, Richard D;  
 APPLICANT: Kuo, Mei-Chang;  
 APPLICANT: Yeung, Siu-mei H.;  
 APPLICANT: Brauer, Andrew;  
 APPLICANT: Exley, Mark A.;  
 APPLICANT: Powers, Steven P.  
 TITLE OF INVENTION: Allergenic Proteins And Peptides From  
 TITLE OF INVENTION: Japanese Cedar Pollen  
 NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.  
 STREET: 610 Lincoln St  
 CITY: Waltham  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02154

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/467,023  
 FILING DATE: June 6, 1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/350,225  
 FILING DATE: December 6, 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872  
 REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 188:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 127 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 FRAGMENT TYPE: internal  
 US-08-467-023-188

Query Match 13.9%; Score 99; DB 3; Length 127;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQNTITG 49  
 DB 83 IDIFASKNFHLQNTITG 101



## RESULT 9

US-08-290-448A-72

Sequence 72, Application US/08290448A

Patent No. 5676954

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce

APPLICANT: Klapper, David G.

APPLICANT: Rafnar, Thorun

APPLICANT: Kuo, Mei-chang

TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,448A

FILING DATE: August 15, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/529,951

FILING DATE: May 29, 1990

APPLICATION NUMBER: US 07/325,365

FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: IMI-018CN

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 72:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-290-448A-72

Query Match 13.9%; Score 98.5; DB 1; Length 387;

Best Local Similarity 34.7%; Pred. No. 0.00012;

Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIFSGNNMNIKMKMPYIAGYTFDGRRAEVSVAH---VNGAK--FIRVD----- 121

DB 104 RPLMIFERDWMVIRLDKEMVNSDKTIDRGAKVEIINAGFTLNGVKNVLIHINMMDVK 163

QY 122 ---GIIAAYQNPAAS 132

DB 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,448A

FILING DATE: August 15, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/529,951

FILING DATE: May 29, 1990

APPLICATION NUMBER: US 07/325,365

FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragouras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: IMI-018CN

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 72:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-290-448A-72

Query Match 13.9%; Score 98.5; DB 1; Length 387;

Best Local Similarity 34.7%; Pred. No. 0.00012;

Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIFSGNNMNIKMKMPYIAGYTFDGRRAEVSVAH---VNGAK--FIRVD----- 121

DB 104 RPLMIFERDWMVIRLDKEMVNSDKTIDRGAKVEIINAGFTLNGVKNVLIHINMMDVK 163

QY 122 ---GIIAAYQNPAAS 132

DB 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

Db 164 VNPGLIKSNDGPAA 178

FILING DATE: May 29, 1990  
 APPLICATION NUMBER: US 07/325,365  
 FILING DATE: March 17, 1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amy E. Mandragouras  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: IMI-018DV  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 TELEFAX: (617)227-5941  
 INFORMATION FOR SEQ ID NO: 72:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 387 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-175-069A-72

Query Match 13.9%; Score 98.5; DB 1; Length 387;  
 Best Local Similarity 34.7%; Pred. No. 0.00012;  
 Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

Oy 73 RPLMTIFSGNNMIKIKMPMTAGYKTFDGRRAEVSYYH---VNGAK--FIRRD----- 121  
 Db 104 RPLMTIFSRDVMIRLDKEMVNSDKTIDRGAKVEIINAGFTLNGVXVLIHNNHMDVK 163  
 Oy 122 ---GIIAAYONPAS 132  
 Db 164 VNPGLIKSNDGPAA 178

RESULT 12  
 US-08-461-939B-72  
 Sequence 72, Application US/08461939B  
 Patent No. 6335019

GENERAL INFORMATION:  
 APPLICANT: Rogers, Bruce  
 APPLICANT: Klapper, David G.  
 APPLICANT: Rafnar, Thorunn  
 APPLICANT: Kuo, Mei-chang  
 TITLE OF INVENTION: Methods For Treating Sensitivity To A  
 TITLE OF INVENTION: Protein Allergen Using Peptides Which Include A T Cell Epitope  
 NUMBER OF SEQUENCES: 93  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD, LLP  
 STREET: 28 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/461,939B  
 FILING DATE:

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/464,000  
 FILING DATE: 05-JUN-1995  
 APPLICATION NUMBER: US 08/290,448  
 FILING DATE: 15-AUG-1994  
 APPLICATION NUMBER: US 07/529,951  
 FILING DATE: 29-MAY-1990  
 APPLICATION NUMBER: US 07/325,365  
 FILING DATE: 17-MAR-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amy E. Mandragouras  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: IMI-018CNDV  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400

TELEFAX: (617)742-4214  
 INFORMATION FOR SEQ ID NO: 72:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 387 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-461-939B-72

Query Match 13.9%; Score 98.5; DB 4; Length 387;  
 Best Local Similarity 34.7%; Pred. No. 0.00012;  
 Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

Oy 73 RPLMTIFSGNNMIKIKMPMTAGYKTFDGRRAEVSYYH---VNGAK--FIRRD----- 121  
 Db 104 RPLMTIFSRDVMIRLDKEMVNSDKTIDRGAKVEIINAGFTLNGVXVLIHNNHMDVK 163  
 Oy 122 ---GIIAAYONPAS 132  
 Db 164 VNPGLIKSNDGPAA 178

RESULT 13  
 US-08-464-000-72  
 Sequence 72, Application US/08464000  
 Patent No. 6335020

GENERAL INFORMATION:  
 APPLICANT: Rogers, Bruce  
 APPLICANT: Klapper, David G.  
 APPLICANT: Rafnar, Thorunn  
 APPLICANT: Kuo, Mei-chang  
 TITLE OF INVENTION: Allergenic Peptides from Ragweed Pollen  
 NUMBER OF SEQUENCES: 93  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD, LLP  
 STREET: 60 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/464,000  
 FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/290,448  
 FILING DATE: 15-AUG-1994  
 APPLICATION NUMBER: US 07/529,951  
 FILING DATE: 29-MAY-1990  
 APPLICATION NUMBER: US 07/325,365  
 FILING DATE: 17-MAR-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Amy E. Mandragouras  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: IMI-018CND2

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617)227-7400  
 TELEFAX: (617)227-5941  
 INFORMATION FOR SEQ ID NO: 72:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 387 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-464-000-72

Query Match 13.9%; Score 98.5; DB 4; Length 387;  
 Best Local Similarity 34.7%; Pred. No. 0.00012;  
 Matches 26; Conservative 12; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIISGNNIKLKMPYIAGYKTFDGRRAEVSYYH-----VNGAK--FIRVD----- 121  
DB 104 RPLMIIFERDVIRLDKEMVYVNSDKTIDGARGAKVEIINAGFTLNGKNTIHNINHDYK 163  
QY 122 -----GIYAAYONPAS 132  
DB 164 VNBGLIKSNDGPAA 178

## RESULT 14

US-08-290-448A-76  
Sequence 76, Application US/08290448A  
Patent No. 5676954

## GENERAL INFORMATION:

APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Ratnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,448A  
FILING DATE: August 15, 1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941

## INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-290-448A-76

## Query Match

Best Local Similarity 13.6%; Score 97; DB 1; Length 397;  
Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;

QY 52 WKNNRIWLO-----FAKLT-----GFTLMGRRPLMIIF 79  
DB 61 WENNROALADCAQGFRAKTYGGKMGDVYVTSNDDVDVANPEKGTLPFAAQNRPFLMIIF 120  
QY 80 SGNNIKLKMPYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124  
DB 121 KNDVYININQELVNSDKTIDRGYKVEI--INGGLTLMVKNII 163

RESULT 15  
US-08-290-448A-76  
Sequence 76, Application US/08290448A  
Patent No. 5698204

GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Ratnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875

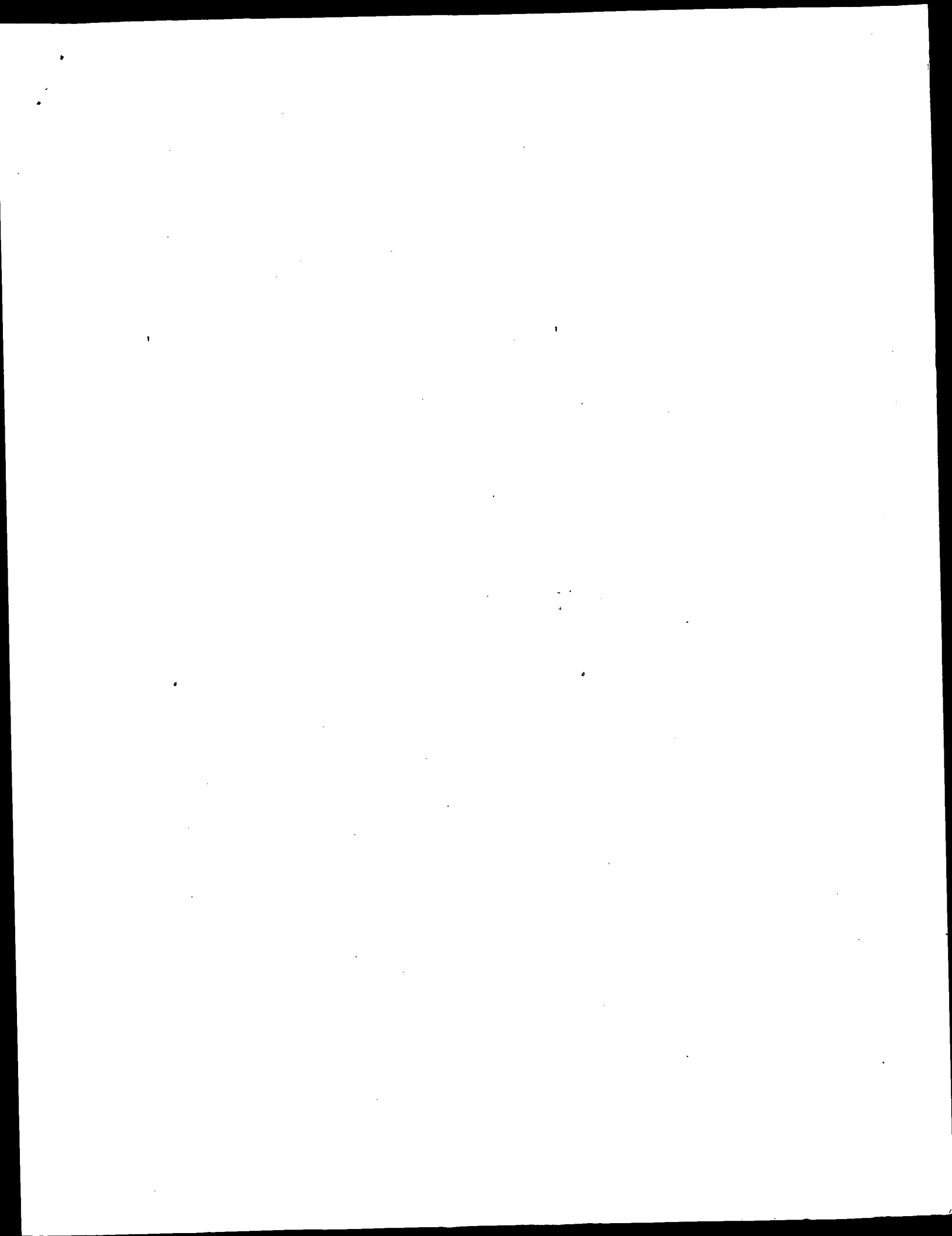
## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,448A  
FILING DATE: August 15, 1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-290-448A-76

Query Match 13.6%; Score 97; DB 1; Length 397;  
Best Local Similarity 29.5%; Pred. No. 0.0002;  
Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;

QY 52 WKNNRIWLO-----FAKLT-----GFTLMGRRPLMIIF 79  
DB 61 WENNROALADCAQGFRAKTYGGKMGDVYVTSNDDVDVANPEKGTLPFAAQNRPFLMIIF 120  
QY 80 SGNNIKLKMPYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124  
DB 121 KNDVYININQELVNSDKTIDRGYKVEI--INGGLTLMVKNII 163

Search completed: April 23, 2003, 18:05:34  
Job time: 17 secs



GenCore version 5.1.4 ps 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 / Search time 67.5789 seconds  
(without alignments)  
243.919 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406  
Sequence: 1 MKVTVAFNQGPNRRVFIK.....IASRRVDGIIAAYQNPASWK 80

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL.21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriophage.\*  
17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	18.0	367	10 Q9L1T2	Q9L1T2 juniperus v
2	73	18.0	367	10 Q9L1T1	Q9L1T1 juniperus v
3	73	18.0	367	10 Q9M4S6	Q9M4S6 cupressus s
4	73	18.0	367	10 Q9M4S5	Q9M4S5 cupressus s
5	73	18.0	367	10 Q9M4S4	Q9M4S4 cupressus s
6	73	18.0	367	10 Q9M4S3	Q9M4S3 cupressus s
7	73	18.0	367	10 Q9M4S2	Q9M4S2 cupressus s
8	73	18.0	367	10 Q9M4S1	Q9M4S1 juniperus o
9	73	18.0	367	10 Q93XL6	Q93XL6 cupressus a
10	71	17.5	374	10 Q9RUB1	Q9RUB1 juniperus a
11	71	17.5	507	10 Q9FY19	Q9FY19 cryptomeria
12	70	17.2	152	13 Q90Y05	Q90Y05 juniperus a
13	69.5	17.1	335	2 Q9AHF2	Q9AHF2 ictalurus p
14	68	16.7	394	10 Q654S6	Q654S6 arabidopsis
15	67	16.5	674	16 Q9KD03	Q9KD03 bacillus ha
16	66.5	16.4	155	2 Q93S55	Q93S55 mycobacteri

17	66	16.3	350	16 P73865	P73865 synchocyst
18	66	16.3	394	10 Q654S7	Q654S7 arabidopsis
19	66	16.3	464	16 Q99W16	Q99W16 staphylococ
20	65.5	16.1	391	17 Q9YB91	Q9YB91 aeropyrum p
21	65.5	16.1	531	5 Q9TX24	Q9TX24 caenorhabdi
22	65.5	16.1	1758	16 Q9QMS5	Q9QMS5 escherichia
23	65	16.0	114	2 Q9X790	Q9X790 mycobacteri
24	65	16.0	132	5 Q962R1	Q962R1 spidoptera
25	65	16.0	300	10 Q65388	Q65388 arabidopsis
26	65	16.0	488	12 Q88525	Q88525 turkey herp
27	64.5	15.9	770	16 Q8YL18	Q8YL18 anabana sp
28	64	15.8	232	8 Q21280	Q21280 reclinomona
29	64	15.8	398	16 Q9WZL9	Q9WZL9 thermotoga
30	64	15.8	433	16 Q9K093	Q9K093 vibrio chol
31	63.5	15.6	533	5 Q01617	Q01617 caenorhabdi
32	63	15.5	645	16 Q98504	Q98504 rhizobium 1
33	63	15.5	753	13 Q9DE07	Q9DE07 gallus gall
34	62.5	15.4	629	4 Q9UPY1	Q9UPY1 homo sapien
35	62	15.3	244	16 Q9A100	Q9A100 streptococ
36	62	15.3	368	17 Q97Z14	Q97Z14 sulfobolus
37	62	15.3	559	4 Q9H796	Q9H796 homo sapien
38	62	15.3	567	16 Q91120	Q91120 pseudomonas
39	61.5	15.1	102	2 Q66004	Q66004 mycoplasma
40	61.5	15.1	228	17 Q977C1	Q977C1 sulfobolus
41	61.5	15.1	369	16 Q9RZU0	Q9RZU0 deinococcus
42	61.5	15.1	543	10 Q9PMV3	Q9PMV3 arabidopsis
43	61.5	15.1	730	2 Q9X571	Q9X571 rhizobium 1
44	61	15.0	142	5 Q9GPA9	Q9GPA9 chetax deat
45	61	15.0	173	2 Q47446	Q47446 escherichia

## ALIGNMENTS

RESULT 1  
ID Q9L1T2 PRELIMINARY; PRT; 367 AA.

AC 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Pollen major allergen 1-2.  
OS Juniperus virginiana (Eastern red cedar).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferales; Coniferales; Cupressaceae; Juniperus.  
OX NCBI\_TaxId=39584;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21315424; PubMed=11422137;  
RA "Midoro-Horuti T.M., Goldblum R.M., Brooks E.G.;  
RT "Identification of mutations in the genes for the pollen allergens of  
RT eastern red cedar (Juniperus virginiana).";  
RL Clin. Exp. Allergy 31:771-778 (2001).  
DR EMBL; AF151427; AAFA80164.1; -  
DR InterPro; IPR002022; Amb allergen.  
DR Pfam; PF00544; pec lyase; 1.  
DR PRINTS; PRO0807; AMBALRGEN.  
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
Best Local Similarity 55.6%; Pred. No. 2.3;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAFNQGPNRRVFIKRVSVITH 27  
Db 234 MKVTVAFNQGPNAGGPMRPRARYGLVH 260

RESULT 2  
ID Q9L1T1 PRELIMINARY; PRT; 367 AA.  
AC Q9L1T1;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Pollen major allergen 1-1.  
 OS Juniperus virginiana (Eastern red cedar).  
 OC Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 OK NCBI\_TaxID=39584;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21315424; PubMed=11422137;  
 RA Midoro-Horiuti T.M., Goldblum R.M., Brooks E.G.;  
 RT "Identification of mutations in the genes for the pollen allergens of  
 eastern red cedar (Juniperus virginiana).";  
 RL Clin. Exp. Allergy 31:771-778(2001).  
 DR EMBL; AF151429; AAF80166.1; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec\_lyase1.1.  
 DR PRINTS; PR00807; AMBAALLERGEN.  
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E711 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
 Best Local Similarity 55.6%; Pred. No. 2.3;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
 DB 234 MKVTAFNQGPNAGQMPRPRARYGLVH 260

RESULT 3

ID Q9M4S6 PRELIMINARY; PRT; 367 AA.  
 AC Q9M4S6;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Cup s 1 pollen allergen.  
 GN CUP81.  
 OS Cupressus sempervirens.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 OK NCBI\_TaxID=13469;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Monsalve R.I., Villalba M., Rodriguez R.;  
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen  
 of Cupressus sempervirens.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF257491; AAF72625.1; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec\_lyase1.1.  
 DR PRINTS; PR00807; AMBAALLERGEN.  
 SQ SEQUENCE 367 AA; 39835 MW; B4B9C60108C2C5A3 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
 Best Local Similarity 55.6%; Pred. No. 2.3;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
 DB 234 MKVTAFNQGPNAGQMPRPRARYGLVH 260

RESULT 4  
 ID Q9M4S5 PRELIMINARY; PRT; 367 AA.  
 AC Q9M4S5;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Cup s 1 pollen allergen.  
 GN CUP81.  
 OS Cupressus sempervirens.

OC Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 OK NCBI\_TaxID=13469;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Monsalve R.I., Villalba M., Rodriguez R.;  
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen  
 of Cupressus sempervirens.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF257492; AAF72626.1; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec\_lyase1.1.  
 DR PRINTS; PR00807; AMBAALLERGEN.  
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC03263B741 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
 Best Local Similarity 55.6%; Pred. No. 2.3;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
 DB 234 MKVTAFNQGPNAGQMPRPRARYGLVH 260

RESULT 5

ID Q9M4S4 PRELIMINARY; PRT; 367 AA.  
 AC Q9M4S4;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Cup s 1 pollen allergen.  
 GN CUP81.  
 OS Cupressus sempervirens.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 OK NCBI\_TaxID=13469;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Monsalve R.I., Villalba M., Rodriguez R.;  
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen  
 of Cupressus sempervirens.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF257493; AAF72627.1; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec\_lyase1.1.  
 DR PRINTS; PR00807; AMBAALLERGEN.  
 SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
 Best Local Similarity 55.6%; Pred. No. 2.3;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
 DB 234 MKVTAFNQGPNAGQMPRPRARYGLVH 260

RESULT 6  
 ID Q9M4S3 PRELIMINARY; PRT; 367 AA.  
 AC Q9M4S3;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Cup s 1 pollen allergen.  
 GN CUP81.  
 OS Cupressus sempervirens.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 OK NCBI\_TaxID=13469;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Monsalve R.I., Villalba M., Rodriguez R.;  
RT "Cloning and expression of Cup s 1, the major allergen of the pollen  
of Cupressus sempervirens.";  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF257494; AAF72628.1; -  
DR InterPro: IPR002022; Amb\_allergen.  
DR Pfam: PF00544; pec\_lyase.1.  
DR PRINTS: PR00807; AMBALLERGEN.  
SQ SEQUENCE 367 AA; 39832 MW; B5DFB5A61C07A53 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
Best Local Similarity 55.6%; Pred. No. 2.3;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 7  
ID Q9M4S2 PRELIMINARY; PRT; 367 AA.  
AC Q9M4S2;  
DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE Cup s 1 pollen allergen.  
GN CUPSI.  
OS Cupressus sempervirens.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
OX NCBI\_TaxID=13469;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Monsalve R.I., Villalba M., Rodriguez R.;  
RT "Cloning and expression of Cup s 1, the major allergen of the pollen  
of Cupressus sempervirens.";  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF257495; AAF72629.1; -  
DR InterPro: IPR002022; Amb\_allergen.  
DR Pfam: PF00544; pec\_lyase.1.  
DR PRINTS: PR00807; AMBALLERGEN.  
SQ SEQUENCE 367 AA; 39819 MW; AF7E055A61C07A53 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
Best Local Similarity 55.6%; Pred. No. 2.3;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 8  
ID Q93X51 PRELIMINARY; PRT; 367 AA.  
AC Q93X51;  
DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)  
DE Putative allergen jun o 1.  
GN JUN O 1.  
OS Juniperus oxycedrus (Prickly juniper).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
OX NCBI\_TaxID=69008;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE=POLLEN;  
RA Iacovacci P., Di Felice G., Pini C.;  
RT "Cloning of Juniperus oxycedrus major allergen.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ293767; CAC48400.1; -

DR InterPro: IPR002022; Amb\_allergen.  
DR Pfam: PF00544; pec\_lyase.1.  
SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFDB1B9D7 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
Best Local Similarity 55.6%; Pred. No. 2.3;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 9  
ID Q93XL6 PRELIMINARY; PRT; 367 AA.  
AC Q93XL6;  
DT 01-DEC-2001 (T-EMBLrel. 19, Created)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE Putative allergen Cup a 1 precursor.  
GN CUP A 1.  
OS Cupressus arizonica.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
OX NCBI\_TaxID=49011;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Butters R., Di Felice G., Pini C.;  
RT "Cloning of Cupressus arizonica major allergen.";  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ278498; CAC37790.2; -  
KM Signal.  
FT SIGNAL. 1 21 POTENTIAL.  
FT CHAIN 22 367 PUTATIVE ALLERGEN CUP A 1.  
SQ SEQUENCE 367 AA; 39809 MW; AFE97260423A9F28 CRC64;

Query Match 18.0%; Score 73; DB 10; Length 367;  
Best Local Similarity 55.6%; Pred. No. 2.3;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNQGPNRRVFIKRVSNVITH 27  
DB 234 MKVTAFNQGPNRRVFIKRVSNVITH 260

RESULT 10  
ID Q8RUR1 PRELIMINARY; PRT; 374 AA.  
AC Q8RUR1;  
DT 01-JUN-2002 (T-EMBLrel. 21, Created)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE CRY J 1 precursor.  
GN CRY J 1.1 OR CRY J 1.2.  
OS Cryptomeria japonica (Japanese cedar).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.  
OX NCBI\_TaxID=3369;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE=POLLEN;  
RA Futamura N., Shinohara K.;  
RT "Isolation and characterization of cDNAs encoding major allergen Cry j  
1 from Cryptomeria japonica pollen.";  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB081309; BAB86286.1; -  
DR EMBL: AB081310; BAB86287.1; -  
KM Signal.  
FT SIGNAL. 1 21 POTENTIAL.  
FT CHAIN 22 374 CRY J 1.  
SQ SEQUENCE 374 AA; 40720 MW; 90D0085D24BF2BD4 CRC64;

Query Match 17.5%; Score 71; DB 10; Length 374;  
 Best Local Similarity 55.6%; Pred. No. 4.1;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTAFNPGFNNRRVFIKRVSNVTH 27  
 |||||  
 DB 234 MKVTAFNPGFNNRRVFIKRVSNVTH 260

RESULT 11

Q9FY19 PRELIMINARY; PRT; 507 AA.

AC Q9FY19; 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Pollen major allergen 2 protein precursor.  
 GN JNA2.  
 OS Juniperus ashei (Ozark white cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 NX NCBI\_TaxID=13101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MALE POLLEN;  
 RX MEDLINE=20403896; PubMed=10944464;  
 RA Yoshizawa M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;  
 RT "Purification, identification and cDNA cloning of Jun a 2, the second  
 major allergen of mountain cedar pollen."  
 RL Biochem. Biophys. Res. Commun. 275:195-202(2000).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES  
 (POLYGALACTURONASES).  
 CC EMBL: AJ404653; CAC05582.1; -  
 DR HSPB; P26509; 188.  
 DR InterPro; IPR000743; GH28.  
 DR InterPro; IPR000408; Reg\_chir\_condens.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN\_1.  
 DR PROSITE; PS00626; RCCL\_2; UNKNOWN\_1.  
 KW Cell wall; Glycosidase; Hydrolase; Signal.  
 FT SIGNAL 1 54  
 SQ SEQUENCE 507 AA; 55730 MW; 28B0AA5B58FE5A CRC64;

Query Match 17.5%; Score 71; DB 10; Length 507;  
 Best Local Similarity 73.7%; Pred. No. 5.9;  
 Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49  
 |||||  
 DB 237 IDIFASKNFHLQKNTIGTG 255

RESULT 12

Q90Y05 PRELIMINARY; PRT; 152 AA.

AC Q90Y05; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE 40S ribosomal protein S18.  
 OS Ictalurus punctatus (Channel catfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;  
 OC Ictaluridae; Ictalurus.  
 NX NCBI\_TaxID=7998;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Karsl A., Patterson A., Feng J., Liu Z.J.;  
 RT "Translational machinery of channel catfish: I. A transcriptomic  
 approach to the analysis of 32 40S ribosomal protein genes and their  
 expression";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF402827; AAK95201.1; -  
 DR InterPro; IPR001892; Ribosomal\_S13.  
 DR Pfam; PF00416; Ribosomal\_S13; 1.  
 DR ProDom; PD001363; Ribosomal\_S13; 1.  
 DR PROSITE; PS00646; RIBOSOMAL\_S13; UNKNOWN\_1.  
 KW Ribosomal protein.  
 SQ SEQUENCE 152 AA; 17691 MW; 1F018266809D92A1 CRC64;

Query Match 17.2%; Score 70; DB 13; Length 152;  
 Best Local Similarity 29.1%; Pred. No. 1.9;  
 Matches 23; Conservative 8; Mismatches 32; Indels 16; Gaps 2;

QY 12 ENRRVFIRGVNVIHGRIDIFASKNFHLQKNTIGTGR-----ISLTKTSKGI 61  
 |||||  
 DB 6 ENRRVFIRGVNVIHGRIDIFASKNFHLQKNTIGTGR-----ISLTKTSKGI 59

QY 62 ASRRVDGIITAAVONPASKW 80  
 |||||  
 DB 60 TEDEVRRVVTIMQNPQYK 78

RESULT 13

Q9AHP2 PRELIMINARY; PRT; 335 AA.

AC Q9AHP2; 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical 36.5 kDa protein.  
 OS Agrobacterium tumefaciens.  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 NX NCBI\_TaxID=358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D3;  
 RX MEDLINE=21322708; PubMed=11429459;  
 RA Trost S., Bauer R., Knackmuss H.J., Stolz A.;  
 RT "Genetic and biochemical characterization of an enantioselective  
 amidase from Agrobacterium tumefaciens strain d3."  
 RL Microbiology 147:1815-1824(2001).  
 DR EMBL; AF315580; AAK28494.1; -  
 DR InterPro; IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 335 AA; 36505 MW; 334C9B65A972C0C1 CRC64;

Query Match 17.1%; Score 69.5; DB 2; Length 335;  
 Best Local Similarity 38.6%; Pred. No. 5.4;  
 Matches 27; Conservative 4; Mismatches 28; Indels 11; Gaps 3;

QY 10 FGNRRVFIRGVNVIHGRIDIFASKNFHLQKNTIGT-----GRISLTKTSKGI 62  
 |||||  
 DB 240 FGNRRVFIRGVNVIHGRIDIFASKNFHLQKNTIGT-----GRISLTKTSKGI 297

QY 63 SRRVDGIITAA 72  
 |||||  
 DB 298 I--VSGIATA 305

RESULT 14

Q65456 PRELIMINARY; PRT; 394 AA.

AC Q65456; 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Pectate lyase like protein.  
 GN FIN20.180 OR AT4G22080.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.



OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Wedler H., Wamburt R., Bancroft I., Mewes H.W., Mayer K.,  
 RA Scheller C.;  
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Wedler H., Wamburt R., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL022140; CAA1811.1; -  
 DR EMBL; AL161556; CAB79163.1; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR InterPro; IPR002048; EF-Hand.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PRO0807; AMBALERGEN.  
 DR PROSITE; PS0018; EF\_HAND; UNKNOWN\_1.  
 KW Lyase.  
 SQ SEQUENCE 394 AA; 43476 MW; A48D586ECF148CA CRC64;

Query Match 16.7%; Score 68; DB 10; Length 394;  
 Best Local Similarity 28.6%; Pred. No. 9.8;  
 Matches 24; Conservative 7; Mismatches 27; Indels 26; Gaps 3;  
 OY 1 MKVTAFNPGPNRRVFIKRVSVIHH--GRID-----IFASKNPHLOKN 44  
 DB 256 MKVTAFNPGPNRRVFIKRVSVIHH--GRID-----IFASKNPHLOKN 44  
 OY 45 TIGTRRISLTKTSGKIASRRVDG 68  
 DB 316 -----KSNKVTKEKRVG 329

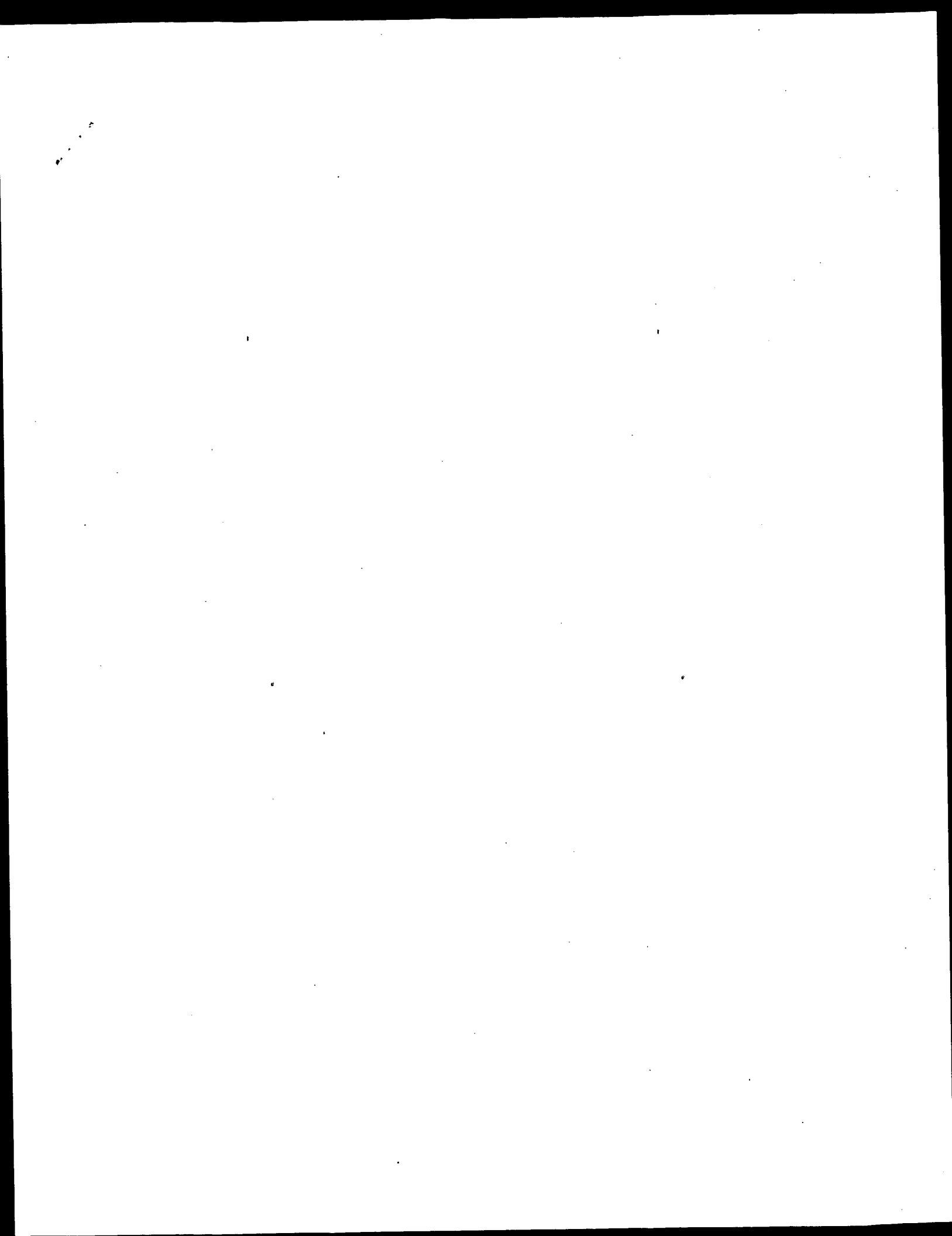
RESULT 15  
 ID 09KD03 PRELIMINARY; PRT; 674 AA.  
 AC 09KD03;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE ABC transporter (ATP-binding protein).  
 GN BH1158.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 OC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-125 / JCM 9153;  
 RX MEDLINE=20512582; PubMed=11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maui N.,  
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kohara S.,  
 RA Horiuchi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 DR EMBL; AP001511; BAB04877.1; -  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR001140; ABCtransportTM.  
 DR InterPro; IPR003439; ABC transporter.  
 DR Pfam; PF00664; ABC membrane; 1.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR ProDom; PD000006; ABC\_transport; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; UNKNOWN\_1.  
 KW ATP-binding; Complete proteome.  
 SQ SEQUENCE 674 AA; 75735 MW; 6AF94B058C84C957 CRC64;

Query Match

16.5%; Score 67; DB 16; Length 674;

Best Local Similarity 29.7%; Pred. No. 24;  
 Matches 19; Conservative 15; Mismatches 20; Indels 10; Gaps 3;  
 OY 1 MKVTAFNPGPNRRVFIKRVSVIHH--GRID-----IFASKNPHLOKN 44  
 DB 158 LMVVASFFQYQ--ORFYLQKAAVRITQRLIDLFN-----HLSRLPV---RFDMMPAOK 207  
 OY 61 IASR 64  
 DB 208 VVSR 211

Search completed: April 20, 2003, 13:12:54  
 Job time : 71.5789 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 97.0526 Seconds  
(without alignments)  
109.838 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406

Sequence: 1 MKYVAENQFGNRRVFIK.....IASRVYDIIAQNPNASWK 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	406	100.0	80	AAW27369	Multi-epitope pept
2	327.5	80.7	105	AAW27370	Multi-epitope pept
3	320	78.8	134	AAW27371	Multi-epitope pept
4	174.5	43.0	214	AAW69120	Cedar pollen aller
5	172	42.4	186	AAW69102	Cedar pollen aller
6	164	40.4	210	AAW68103	Cedar pollen aller
7	144	35.5	47	AAW80353	Sugi allergen prot
8	140.5	34.6	74	AAW23905	Artificial sequenc
9	140.5	34.6	94	AAW23906	Artificial sequenc
10	140.5	34.6	95	AAW23897	Artificial sequenc

11	140.5	34.6	95	21	AAW23901	Artificial sequenc
12	140.5	34.6	95	22	AAW69098	Cedar pollen aller
13	140.5	34.6	96	22	AAW69104	Cedar pollen aller
14	140.5	34.6	96	22	AAW69119	Cedar pollen aller
15	140.5	34.6	97	22	AAW69105	Artificial sequenc
16	140.5	34.6	99	21	AAW23902	Artificial sequenc
17	140.5	34.6	99	21	AAW23907	Artificial sequenc
18	140.5	34.6	99	21	AAW69099	Cedar pollen aller
19	135	33.3	81	19	AAW80358	Sugi allergen prot
20	135	33.3	81	21	AAW23895	Artificial sequenc
21	135	33.3	81	22	AAW69118	Cedar pollen aller
22	135	33.3	88	22	AAW69114	Cedar pollen aller
23	135	33.3	93	21	AAW23881	Artificial sequenc
24	135	33.3	93	21	AAW23896	Artificial sequenc
25	135	33.3	93	22	AAW69097	Cedar pollen aller
26	135	33.3	93	22	AAW69117	Cedar pollen aller
27	133	32.8	71	21	AAW23878	Artificial sequenc
28	133	32.8	81	21	AAW23874	Artificial sequenc
29	133	32.8	81	21	AAW23879	Artificial sequenc
30	133	32.8	81	22	AAW69094	Artificial sequenc
31	127.5	31.4	79	21	AAW23880	Cedar pollen aller
32	127.5	31.4	79	21	AAW23884	Artificial sequenc
33	127.5	31.4	79	22	AAW69095	Cedar pollen aller
34	125.5	30.9	47	19	AAW80352	Sugi allergen prot
35	125.5	30.9	79	21	AAW23885	Artificial sequenc
36	125.5	30.9	79	21	AAW23880	Artificial sequenc
37	125.5	30.9	79	22	AAW69096	Cedar pollen aller
38	121.5	29.9	47	19	AAW80351	Sugi allergen prot
39	120	29.6	47	19	AAW80357	Immunomodulatory p
40	114	28.1	47	19	AAW80356	Artificial sequenc
41	103	25.4	42	22	AAW84105	Artificial sequenc
42	102	25.1	61	21	AAW23894	Artificial sequenc
43	100.5	24.8	61	21	AAW23883	Artificial sequenc
44	99	24.4	19	22	AAW84115	Amino acid sequenc
45	99	24.4	460	16	AAW69791	Japanicum allergen

## ALIGNMENTS

RESULT 1  
AAW27369 standard; peptide; 80 AA.

XX  
AC AAW27369;  
XX  
DT 24-MAR-1998 (first entry)  
XX  
DB Multi-epitope peptide used as immunotherapeutic agent #1.  
XX  
KM Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
XX  
OS Synthetic.  
XX  
PN WO9732600-A1.  
XX  
PD 12-SEP-1997.  
XX  
PF 10-MAR-1997; 97WO-0P00740.  
XX  
PR 10-MAR-1996; 96JP-0080702.  
XX  
PA (MEIP ) MELI MILK PROD CO LTD.  
XX  
PI Daiiriki K, Iwama A, Kino K, Kume A, Sone T;  
XX  
XX WPI; 1997-470495/43.  
XX  
PT Peptide immuno-therapeutic agent to treat allergic diseases -  
PT contains multi-epitope peptide containing T cell epitope regions  
PT from different allergens

PS Claim 6; Page 31; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 80 AA;

Query Match 100.0%; Score 406; DB 18; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 2,6e-48;  
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKVTVAFNQFGPNRRVFIRKVSNTIIHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 60  
 DB 1 MKVTVAFNQFGPNRRVFIRKVSNTIIHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 60  
 QY 61 IASRRVDGIIAAYQNPASWK 80  
 DB 61 IASRRVDGIIAAYQNPASWK 80

RESULT 2  
 AAM27370  
 ID AAM27370 standard; peptide; 105 AA.  
 XX AAM27370;  
 AC AAM27370;  
 XX 24-MAR-1998 (first entry)  
 DT 24-MAR-1998 (first entry)  
 XX Multi-epitope peptide used as immunotherapeutic agent #2.  
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
 XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 OS Synthetic.  
 XX WO9732600-A1.  
 PN WO9732600-A1.  
 XX 12-SEP-1997.  
 PD 12-SEP-1997.  
 XX 10-MAR-1997; 97WO-JP00740.  
 PF 10-MAR-1997; 97WO-JP00740.  
 XX 10-MAR-1996; 96JP-0080702.  
 PR 10-MAR-1996; 96JP-0080702.  
 XX (MEIP) MEIJI MILK PROD CO LTD.  
 PA Dai-ichi K, Iwama A, Kino K, Kume A, Sone T;  
 PI Dai-ichi K, Iwama A, Kino K, Kume A, Sone T;  
 XX WPI; 1997-470495/43.  
 DR WPI; 1997-470495/43.  
 XX Peptide immuno-therapeutic agent to treat allergic diseases -  
 PT contains multi-epitope peptide containing T cell epitope regions  
 PT from different allergens  
 XX  
 PS Claim 6; Page 31; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX SQ Sequence 105 AA;

Query Match 80.7%; Score 327.5; DB 18; Length 105;  
 Best Local Similarity 67.6%; Pred. No. 2,7e-37;  
 Matches 71; Conservative 1; Mismatches 8; Indels 25; Gaps 2;

QY 1 MKVTVAFNQFGPNRRVFIRKVSNTIIHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 53  
 DB 1 MKVTVAFNQFGPNRRVFIRKVSNTIIHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 53  
 QY 54 -LKLTSCKIA-----SRRVDGIIAAYQNPASWK 80  
 DB 61 IASRRVDGIIAAYQNPASWK 80

RESULT 3  
 AAM27371  
 ID AAM27371 standard; peptide; 134 AA.  
 XX AAM27371;  
 AC AAM27371;  
 XX 24-MAR-1998 (first entry)  
 DT 24-MAR-1998 (first entry)  
 XX Multi-epitope peptide used as immunotherapeutic agent #3.  
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
 XX T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 OS Synthetic.  
 XX WO9732600-A1.  
 PN WO9732600-A1.  
 XX 12-SEP-1997.  
 PD 12-SEP-1997.  
 XX 10-MAR-1997; 97WO-JP00740.  
 PF 10-MAR-1997; 97WO-JP00740.  
 XX 10-MAR-1996; 96JP-0080702.  
 PR 10-MAR-1996; 96JP-0080702.  
 XX (MEIP) MEIJI MILK PROD CO LTD.  
 PA Dai-ichi K, Iwama A, Kino K, Kume A, Sone T;  
 PI Dai-ichi K, Iwama A, Kino K, Kume A, Sone T;  
 XX WPI; 1997-470495/43.  
 DR WPI; 1997-470495/43.  
 XX Peptide immuno-therapeutic agent to treat allergic diseases -  
 PT contains multi-epitope peptide containing T cell epitope regions  
 PT from different allergens  
 XX  
 PS Claim 6; Page 32; 58pp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 134 AA;

Query Match 78.8%; Score 320; DB 18; Length 134;  
 Best Local Similarity 53.7%; Pred. No. 4e-36;  
 Matches 72; Conservative 2; Mismatches 6; Indels 54; Gaps 2;

QY 1 MKVTVAFNQFGPNRRVFIRKVSNTIIHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 51  
 DB 1 MKVTVAFNQFGPNRRVFIRKVSNTIIHGRRIDIFASKNFHLQKNTIGGRISLKLTSCK 51  
 QY 52 -----ISLKL-----TSCKIASRRV 66

Db 61 FAKTGTLMGRRLPILIFSGNMNLIKMPMYIAGYKTFDGRRAEVSVMHNGAKFTRRV 120  
 QY 67 DGIIAAVONPASM 80  
 Db 121 DGIIAAVONPASM 134

RESULT 4  
 AAB69120  
 ID AAB69120 standard; Protein; 214 AA.

AC AAB69120;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:61.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; anti-sugipollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.

DR N-PSDB; AAF59044.

PT Novel peptide and its use -

PS Example 11; Page 58-59; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (1) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.

SQ Sequence 214 AA;

Query Match 43.0%; Score 174.5; DB 22; Length 214;  
 Best Local Similarity 38.3%; Pred. No. 9.9e-16;

Matches 49; Conservative 0; Mismatches 2; Indels 77; Gaps 4;

QY 1 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTGTGRRISLKTSGK 60  
 Db 30 MKVTVAFNQFGP-----DIFASKNFHLQKN-----KLTSGK 60

QY 61 IAS-----RRVD--GIITAA 72  
 Db 61 IASCLNGLVHANNVNDPSGKYEKGNITKKEAFNVEQFAKLTGTLGRADPRGIITAA 120

QY 73 YONPASM 80  
 Db 121 YONPASM 128

RESULT 5  
 AAB69102

ID AAB69102 standard; Protein; 186 AA.

AC AAB69102;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:10.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; anti-sugipollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.

DR N-PSDB; AAF59011.

PT Novel peptide and its use -

PS Claim 14; Page 36-37; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (1) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.

SQ Sequence 186 AA;

Query Match 42.4%; Score 172; DB 22; Length 186;  
 Best Local Similarity 40.9%; Pred. No. 1.8e-15;

Matches 47; Conservative 0; Mismatches 2; Indels 66; Gaps 4;

QY 1 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTGTGRRISLKTSGK 60  
 Db 29 MKVTVAFNQFGP-----FASKNFHLQKN-----KLTSGK 58

QY 61 IAS-----RRVD--GIITAAVONPASM 79  
 Db 59 IASCLNGLVHANNVNDPSGKYEKGNITKKEAFNVEQFAKLTGTLGRADPRGIITAAVONPASM 113

RESULT 6

ID AAB69103  
 ID AAB69103 standard; Protein; 210 AA.

AC AAB69103;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:12.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; anti-sugipollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP200032769-A.  
 XX  
 PD 28-NOV-2000.  
 XX  
 PF 15-MAR-2000; 2000JP-00717710.  
 XX  
 PR 15-MAR-1999; 99JP-0068316.  
 XX  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.  
 XX  
 DR WPI; 2001-185061/19.  
 DR N-PSDB; AAF59012.  
 XX  
 PT Novel peptide and its use -  
 XX  
 PS Claim 7; Page 39-40; 75pp; Japanese.  
 XX  
 CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.  
 XX  
 SQ Sequence 210 AA;  
 XX  
 Query Match 40.4%; Score 164; DB 22; Length 210;  
 Best Local Similarity 37.0%; Pred. No. 2.7e-14;  
 Matches 47; Conservative 0; Mismatches 2; Indels 78; Gaps 4;  
 QY 1 MKVTAFNPGPRRRVFIRKSVNVIHGRIDIFASKNPHLOKNTIGRRISLKTSGK 60  
 DB 29 MKVTAFNPGP-----FASKNPHLOKNT-----KLTSGK 58  
 QY 61 IAS-----RRVD-GITAA 72  
 DB 59 IASCLNVLGVANNVNDPSGKVEGNIYTKKEAFNVEQFAKLTGFTLMGRADPRGIITAA 118  
 QY 73 YONPASW 79  
 DB 119 YONPASW 125  
 XX  
 RESULT 7  
 AAM80353  
 ID AAM80353 standard; peptide; 47 AA.  
 XX  
 AC AAM80353;  
 XX  
 DT 11-JAN-1999 (first entry)  
 XX  
 DE Sugi allergen protein Cryj1 derived epitope for T cells.  
 XX  
 KM T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;  
 KM sugi-pollinosis; allergic reaction; pollen.  
 XX  
 OS Synthetic.  
 XX  
 PN JPI0259198-A.  
 XX  
 PD 29-SEP-1998.  
 XX  
 PF 22-DEC-1997; 97JP-0353448.  
 XX  
 PR 24-DEC-1996; 96JP-0343441.  
 XX  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.  
 XX

DR WPI; 1998-577037/49.  
 XX  
 PT A linked T cell epitope peptide - used for the treatment of  
 PT sugi-pollinosis  
 XX  
 PS Claim 10; Page 5; 21pp; Japanese.  
 XX  
 CC AAM80339-58 represent epitopes for T cells, derived from the sugi  
 CC allergen proteins Cryj1 (AAM80339-44, AAM80350-53 and AAM80356-58) and  
 CC Cryj2 (AAM80345-49 and AAM80354-55). The peptides are useful for the  
 CC treatment of sugi-pollinosis, an allergic reaction of the body to  
 CC pollen.  
 XX  
 SQ Sequence 47 AA;  
 XX  
 Query Match 35.5%; Score 144; DB 19; Length 47;  
 Best Local Similarity 44.3%; Pred. No. 2.3e-12;  
 Matches 35; Conservative 0; Mismatches 0; Indels 44; Gaps 2;  
 QY 1 MKVTAFNPGPRRRVFIRKSVNVIHGRIDIFASKNPHLOKNTIGRRISLKTSGK 60  
 DB 13 MKVTAFNPGP-----FASKNPHLOKNT----- 35  
 QY 61 IASRRVDGITAAAYONPASW 79  
 DB 36 -----GITAAAYONPASW 47  
 XX  
 RESULT 8  
 AAB23905  
 ID AAB23905 standard; peptide; 74 AA.  
 XX  
 AC AAB23905;  
 XX  
 DT 17-JAN-2001 (first entry)  
 XX  
 DE Artificial sequence designed peptide #26.  
 XX  
 KM Peptide synthesis; chemical synthesis; solid phase synthesis.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200055182-A1.  
 XX  
 PD 21-SEP-2000.  
 XX  
 PF 15-MAR-2000; 2000WO-0P01584.  
 XX  
 PR 15-MAR-1999; 99JP-0067917.  
 XX  
 PA (SANY ) SANKYO CO LTD.  
 XX  
 PI Kawaguchi J, Serizawa N;  
 XX  
 DR WPI; 2000-602106/57.  
 XX  
 PT Chemically synthesizing a peptide by solid phase synthesis by adding  
 PT protective group to inactivate part of alpha amino group of an amino  
 PT acid group -  
 XX  
 PS Example 6; Page 23; 38pp; Japanese.  
 XX  
 CC The present invention describes a method for chemically synthesizing a  
 CC peptide by solid phase synthesis. The method comprises processing a  
 CC carrier resin to which the C-terminal amino acid derivative of the  
 CC target peptide has been preliminarily bonded, is characterised by adding  
 CC a protective group to inactivate a part of alpha-amino groups of an  
 CC amino acid capable of undergoing a peptide extension reaction on the  
 CC carrier resin. The method is useful for synthesizing a peptide by solid  
 CC phase synthesis. Long chain peptides can be readily synthesised in with  
 CC a high success ratio without using the segment condensation method or  
 CC gene manipulations. The present sequence represents an amino acid  
 CC peptide sequence given in an example from the present invention.

XX Sequence 74 AA;  
SQ

Query Match 34.6%; Score 140.5; DB 21; Length 74;  
Best Local Similarity 48.0%; Pred. No. 1.3e-11;  
Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

DB 1 MKVTAFNQGPNRRVFIKRVSVIIGHRRIDIPASKNPHLOKNTIGTRISLKTSGK 60  
6 MKVTAFNQGPNRRVFIKRVSVIIGHRRIDIPASKNPHLOKNTIGTRISLKTSGK 36  
DY 61 IASRRVDGIIAAYON 75  
37 IASCLNGLVHVANN 51

## RESULT 9

AAB23906  
ID AAB23906 standard; peptide; 94 AA.

XX AAB23906;  
AC

XX 17-JAN-2001 (first entry)  
DT

XX Artificial sequence designed peptide #27.  
DE

XX Peptide synthesis; chemical synthesis; solid phase synthesis.  
KW

XX Synthetic.  
OS

XX WO200055182-A1.  
PN

XX 21-SEP-2000.  
PD

XX 15-MAR-2000; 2000WO-JP01584.  
PF

XX 15-MAR-1999; 99JP-0067917.  
PR

XX (SANY) SANKYO CO LTD.  
PA

XX Kawaguchi J, Serizawa N;  
PI

XX WPI; 2000-602106/57.  
DR

XX Chemically synthesizing a peptide by solid phase synthesis by adding  
PT protective group to inactivate part of alpha amino group of an amino  
XX acid group

XX Example 6; Page 24; 38pp; Japanese.  
PS

XX The present invention describes a method for chemically synthesizing a  
CC peptide by solid phase synthesis. The method comprises processing a  
CC carrier resin to which the C-terminal amino acid derivative of the  
CC target peptide has been preliminarily bonded is characterized by adding  
CC a protective group to inactivate a part of alpha-amino groups of an  
CC amino acid capable of undergoing a peptide extension reaction on the  
CC carrier resin. The method is useful for synthesizing a peptide by solid  
CC phase synthesis. Long chain peptides can be readily synthesized in with  
CC a high success ratio without using the segment condensation method or  
CC gene manipulations. The present sequence represents an amino acid  
CC peptide sequence given in an example from the present invention.

XX Sequence 94 AA;  
SQ

Query Match 34.6%; Score 140.5; DB 21; Length 94;  
Best Local Similarity 48.0%; Pred. No. 1.7e-11;  
Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

DY 1 MKVTAFNQGPNRRVFIKRVSVIIGHRRIDIPASKNPHLOKNTIGTRISLKTSGK 60  
26 MKVTAFNQGPNRRVFIKRVSVIIGHRRIDIPASKNPHLOKNTIGTRISLKTSGK 56  
DY 61 IASRRVDGIIAAYON 75

DB 57 IASCLNGLVHVANN 71  
|||::|

## RESULT 10

AAB23897  
ID AAB23897 standard; peptide; 95 AA.

XX AAB23897;  
AC

XX 17-JAN-2001 (first entry)  
DT

XX Artificial sequence designed peptide SEQ ID NO:5.  
DE

XX Peptide synthesis; chemical synthesis; solid phase synthesis.  
KW

XX Synthetic.  
OS

XX WO200055182-A1.  
PN

XX 21-SEP-2000.  
PD

XX 15-MAR-2000; 2000WO-JP01584.  
PF

XX 15-MAR-1999; 99JP-0067917.  
PR

XX (SANY) SANKYO CO LTD.  
PA

XX Kawaguchi J, Serizawa N;  
PI

XX WPI; 2000-602106/57.  
DR

XX Chemically synthesizing a peptide by solid phase synthesis by adding  
PT protective group to inactivate part of alpha amino group of an amino  
XX acid group

XX Example 5; Page 34-35; 38pp; Japanese.  
PS

XX The present invention describes a method for chemically synthesizing a  
CC peptide by solid phase synthesis. The method comprises processing a  
CC carrier resin to which the C-terminal amino acid derivative of the  
CC target peptide has been preliminarily bonded is characterized by adding  
CC a protective group to inactivate a part of alpha-amino groups of an  
CC amino acid capable of undergoing a peptide extension reaction on the  
CC carrier resin. The method is useful for synthesizing a peptide by solid  
CC phase synthesis. Long chain peptides can be readily synthesized in with  
CC a high success ratio without using the segment condensation method or  
CC gene manipulations. The present sequence represents an amino acid  
CC peptide sequence given in an example from the present invention.

XX Sequence 95 AA;  
SQ

Query Match 34.6%; Score 140.5; DB 21; Length 95;  
Best Local Similarity 48.0%; Pred. No. 1.7e-11;  
Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

DY 1 MKVTAFNQGPNRRVFIKRVSVIIGHRRIDIPASKNPHLOKNTIGTRISLKTSGK 60  
27 MKVTAFNQGPNRRVFIKRVSVIIGHRRIDIPASKNPHLOKNTIGTRISLKTSGK 57  
DY 61 IASRRVDGIIAAYON 75  
58 IASCLNGLVHVANN 72

## RESULT 11

AAB23901  
ID AAB23901 standard; peptide; 95 AA.

XX AAB23901;  
AC

XX 17-JAN-2001 (first entry)  
DT

DE Artificial sequence designed peptide #23.  
 XX Peptide synthesis; chemical synthesis; solid phase synthesis.  
 KW Synthetic.  
 OS WO200055182-A1.  
 PN 21-SEP-2000.  
 PD 15-MAR-2000; 2000WO-JP01584.  
 PF 15-MAR-1999; 99JP-0068316.  
 PR (SANY ) SANKYO CO LTD.  
 PA Kawaguchi J, Serizawa N;  
 PI WPI; 2000-602106/57.  
 DR Chemically synthesizing a peptide by solid phase synthesis by adding  
 XX protective group to inactivate part of alpha amino group of an amino  
 PT acid group -  
 PS Example 5; Page 21-22; 38pp; Japanese.  
 CC The present invention describes a method for chemically synthesizing a  
 CC peptide by solid phase synthesis. The method comprises processing a  
 CC carrier resin to which the C-terminal amino acid derivative of the  
 CC target peptide has been preliminarily bonded is characterized by adding  
 CC a protective group to inactivate a part of alpha-amino groups of an  
 CC amino acid capable of undergoing a peptide extension reaction on the  
 CC carrier resin. The method is useful for synthesizing a peptide by solid  
 CC phase synthesis. Long chain peptides can be readily synthesised in with  
 CC a high success ratio without using the segment condensation method or  
 CC gene manipulations. The present sequence represents an amino acid  
 CC peptide sequence given in an example from the present invention.  
 SQ Sequence 95 AA;  
 Query Match 34.6%; Score 140.5; DB 21; Length 95;  
 Best Local Similarity 48.0%; Pred. No. 1.7e-11;  
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;  
 QY 1 MKTVAFNOGPNRRVFIKRVSVIIGHRRIDIFASKNPHLOKNTIGTGRRLSLKLTSGK 60  
 DB 27 MKTVAFNOGPNRRVFIKRVSVIIGHRRIDIFASKNPHLOKNTIGTGRRLSLKLTSGK 57  
 QY 61 IASRRVDGITAAAYON 75  
 DB 58 IASCLNMGVLHVANN 72  
 RESULT 12  
 AAB69098  
 ID AAB69098 standard; Peptide; 95 AA.  
 AC AAB69098;  
 XX 23-APR-2001 (first entry)  
 DT Cedar pollen allergen T cell epitope derived peptide SEQ ID NO:5.  
 DE Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiIgEipollinosis.  
 XX Cryptomeria japonica.  
 OS Synthetic.  
 PN JP2000327699-A.  
 PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.  
 XX 15-MAR-1999; 99JP-0068316.  
 PR (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.  
 PA WPI; 2001-185061/19.  
 DR Novel peptide and its use -  
 XX Claim 5; Page 33-34; 75pp; Japanese.  
 CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an antiIgEipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.  
 SQ Sequence 95 AA;  
 Query Match 34.6%; Score 140.5; DB 22; Length 95;  
 Best Local Similarity 48.0%; Pred. No. 1.7e-11;  
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;  
 QY 1 MKTVAFNOGPNRRVFIKRVSVIIGHRRIDIFASKNPHLOKNTIGTGRRLSLKLTSGK 60  
 DB 27 MKTVAFNOGPNRRVFIKRVSVIIGHRRIDIFASKNPHLOKNTIGTGRRLSLKLTSGK 57  
 QY 61 IASRRVDGITAAAYON 75  
 DB 58 IASCLNMGVLHVANN 72  
 RESULT 13  
 AAB69104  
 ID AAB69104 standard; Protein; 96 AA.  
 AC AAB69104;  
 XX 23-APR-2001 (first entry)  
 DT Cedar pollen allergen T cell epitope derived protein SEQ ID NO:13.  
 DE Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiIgEipollinosis.  
 XX Cryptomeria japonica.  
 OS Synthetic.  
 PN JP2000327699-A.  
 PD 28-NOV-2000.  
 PF 15-MAR-2000; 2000JP-0071710.  
 PR 15-MAR-1999; 99JP-0068316.  
 PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.  
 PA WPI; 2001-185061/19.  
 DR Novel peptide and its use -  
 XX Claim 8; Page 40; 75pp; Japanese.  
 CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:



CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.  
 XX

Sequence 96 AA;

Query Match 34.6%; Score 140.5; DB 22; Length 96;  
 Best Local Similarity 48.0%; Pred. No. 1.8e-11;  
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

OY 1 MKVTVAFNQGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRRISLKLTSCK 60  
 DB 15 MKVTVAFNQGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRRISLKLTSCK 45  
 OY 61 IASRRVDGIIAAQON 75  
 DB 46 IASCLNTGIVHVANN 60

RESULT 14  
 AAB69119  
 ID AAB69119 standard; Protein; 96 AA.  
 AC AAB69119;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:53.  
 XX  
 KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KM T cell epitope; anti-sugipollinosis.  
 XX  
 OS Cryptomeria japonica.  
 OS Synthetic.  
 OS  
 PN JP2000327699-A.  
 XX  
 XX 28-NOV-2000.  
 PD  
 PF 15-MAR-2000; 2000JP-0071710.  
 XX  
 PR 15-MAR-1999; 99JP-0068316.  
 XX  
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY) SANKYO CO LTD.  
 DR WPI; 2001-185061/19.  
 DR N-PSDB; AAF59037.  
 XX  
 PT Novel peptide and its use -  
 XX  
 PS Example 11; Page 54-55; 75pp; Japanese.  
 XX  
 CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (1) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.  
 XX

Sequence 96 AA;

Query Match 34.6%; Score 140.5; DB 22; Length 96;  
 Best Local Similarity 48.0%; Pred. No. 1.8e-11;  
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

OY 1 MKVTVAFNQGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRRISLKLTSCK 60  
 DB 15 MKVTVAFNQGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRRISLKLTSCK 45  
 OY 61 IASRRVDGIIAAQON 75  
 DB 46 IASCLNTGIVHVANN 60

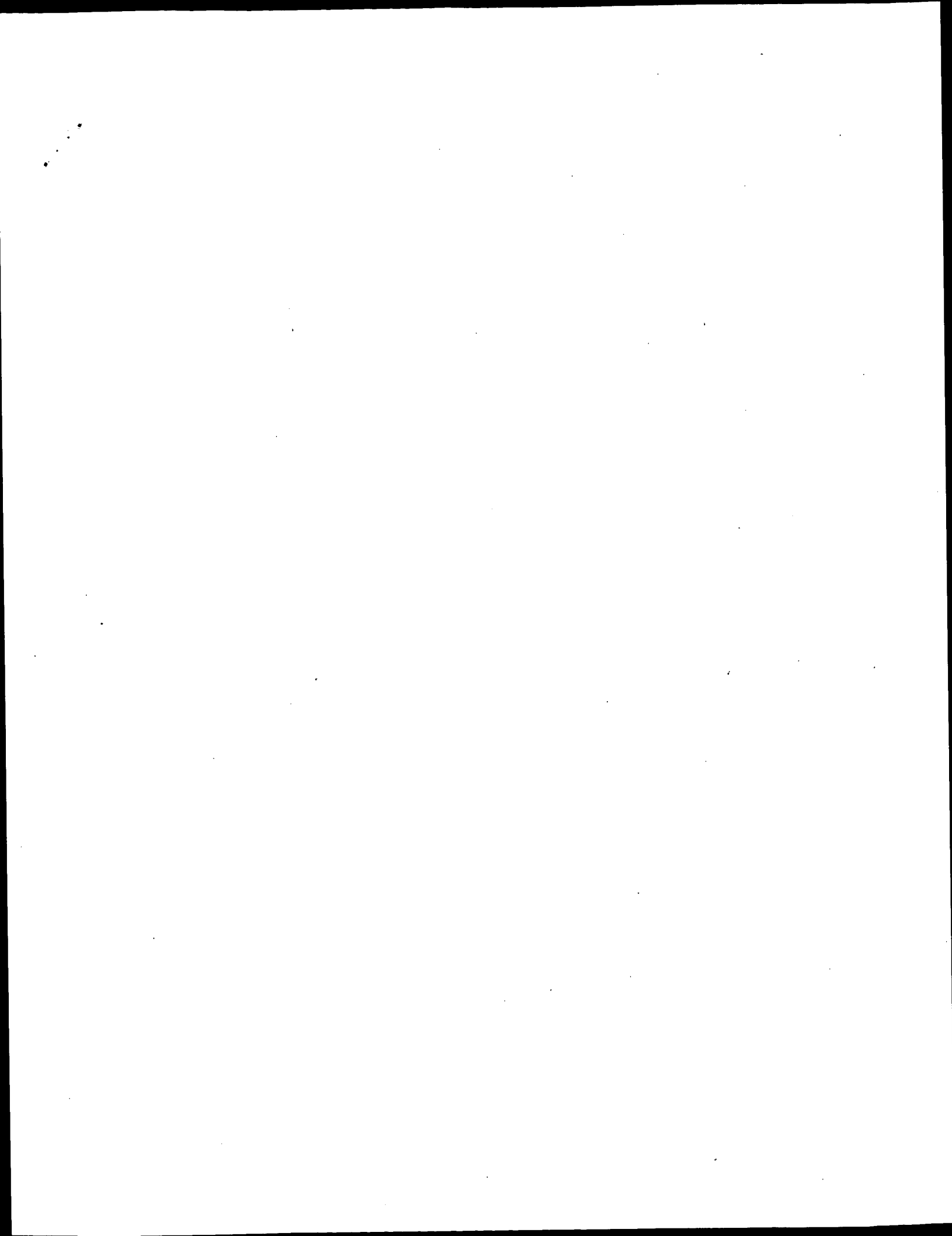
RESULT 15  
 AAB69105  
 ID AAB69105 standard; Protein; 97 AA.  
 AC AAB69105;  
 XX  
 DT 23-APR-2001 (first entry)  
 XX  
 DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:15.  
 XX  
 KW Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KM T cell epitope; anti-sugipollinosis.  
 XX  
 OS Cryptomeria japonica.  
 OS Synthetic.  
 OS  
 PN JP2000327699-A.  
 XX  
 XX 28-NOV-2000.  
 PD  
 PF 15-MAR-2000; 2000JP-0071710.  
 XX  
 PR 15-MAR-1999; 99JP-0068316.  
 XX  
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY) SANKYO CO LTD.  
 DR WPI; 2001-185061/19.  
 DR N-PSDB; AAF59013.  
 XX  
 PT Novel peptide and its use -  
 XX  
 PS Claim 9; Page 41-42; 75pp; Japanese.  
 XX  
 CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (1) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an anti-sugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.  
 XX

Sequence 97 AA;

Query Match 34.6%; Score 140.5; DB 22; Length 97;  
 Best Local Similarity 48.0%; Pred. No. 1.8e-11;  
 Matches 36; Conservative 2; Mismatches 8; Indels 29; Gaps 2;

OY 1 MKVTVAFNQGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRRISLKLTSCK 60  
 DB 16 MKVTVAFNQGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRRISLKLTSCK 46  
 OY 61 IASRRVDGIIAAQON 75  
 DB 47 IASCLNTGIVHVANN 61

Search completed: April 20, 2003, 13:06:02  
 Job time : 98.0526 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 42.5526 Seconds  
(without alignments)  
237.215 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 558  
Sequence: 1 MCVTVAFNQFGPNRRVFIKR.....FDGRVDGIIAAYQNPASWK 105

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 73:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115.5	20.7	514	2	S48730 Cry j II protein - second major aller
2	115	20.6	514	2	JC2498 polygalacturonase
3	98	17.6	514	2	JC7100 major allergen Cry
4	93.5	16.8	374	2	JC2124 major allergen Cry
5	93.5	16.8	374	2	JC2123 Jun a 2 protein -
6	91.5	16.4	507	2	JC7366 F72335
7	84.5	15.1	398	2	F72335 probable synaptochr
8	81	14.5	219	2	C84647 synaptochr
9	77	13.8	219	2	T04630 hypothetical prote
10	75.5	13.5	477	2	S76496 hypothetical prote
11	73.5	13.2	644	2	AD3471 adenyl-yl-sulfate k
12	71	12.7	576	1	AA5049 4-alpha-glucanotri
13	68.5	12.3	390	2	B37154 phosphopentomutase
14	67.5	12.1	1070	2	S46755 hypothetical prote
15	67	12.0	689	1	JW0107 very-long-chain ac
16	66.5	11.9	110	2	D48562 hypothetical nucle
17	66.5	11.9	777	2	T13561 NADH2 dehydrogenas
18	66.5	11.8	777	2	G69310 NADH2 dehydrogenas
19	66	11.8	188	2	B72450 penicillin G acyla
20	66	11.8	196	2	S54580 hypothetical prote
21	65.5	11.7	493	2	AB3032 probable membrane
22	65.5	11.7	602	2	AD1542 hypothetical prote
23	64.5	11.6	317	2	AD1542 exer protein (A722
24	64	11.5	284	2	JC2330 oxidoreductases ho
25	64	11.5	284	2	JC2330 hypothetical 22k p
26	64	11.5	350	2	S75065 major antigenic pr
27	64	11.5	612	2	H81043 sensory transducti
28	64	11.5	612	2	H81043 NADH2 dehydrogenas
29	64	11.5	3194	2	D71917 toxin-like outer m

## ALIGNMENTS

### RESULT 1

S48730  
Cry j II protein - Japanese cedar

C/Species: Cryptomera japonica (Japanese cedar)

C/Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999

C/Accession: S48730

R/Namba, M.; Kurose, M.; ToriGo, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kuri

FEBS Lett. 353, 124-128, 1994

A/Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar p

A/Reference number: S48730, MUID:95010777, PMID:7926035

A/Accession: S48730

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-514 <N>M>

A/Cross-references: GB:D37765; NID:G577695; PIDN:BA07021.1; PID:di007598; PID:G577696

### Query Match

Best Local Similarity 20.7%; Score 115.5; DB 2; Length 514;  
Pred. No. 2.1e-05;  
Matches 29; Conservative 9; Mismatches 11; Indels 21; Gaps 4;

QY	12	PNRRVFIKRVSNVLIHGR-----RID--IFASKKFIHQKNTIGRWKNNRWLOPA 62
DB	96	GNKKKPV--VNNLFNFGPCQPHFTFKVDGIIAAYQN-----PASKNRRWLOPA 143
QY	63	KLTGFTLMGR 72
DB	144	KLTGFTLMGR 153

### RESULT 2

JC2498  
second major allergen Cry j II precursor - Japanese cedar

C/Species: Cryptomera japonica (Japanese cedar)

C/Date: 16-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 21-Jul-2000

C/Accession: JC2498; PC2346; A60147

R/Komiyama, N.; Some, T.; Shimizu, K.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994

A/Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese c

A/Reference number: JC2498; MUID:94271186; PMID:8002972

A/Accession: JC2498

A/Molecule type: mRNA

A/Residues: 1-514 <K>M>

A/Cross-references: DDBJ:D29772; NID:G506857; PIDN:BA06172.1; PID:G506858

A/Accession: PC2346

A/Molecule type: protein

A/Residues: 52-61 <K>O2>

R/Sakaguchi, M.; Inoue, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhai, T.

Allergy 45, 309-312, 1990

A/Title: Identification of the second major allergen of Japanese cedar pollen.

A/Reference number: A60147; MUID:90342988; PMID:2382797

A/Accession: A60147

A/Molecule type: protein

A:Residues: 55-64 <SAK>

C:Keywords: glycoprotein; pollen

F:1-54/Domain: signal sequence #status predicted <SIG>

F:55-460/Product: second major allergen Cry j #status predicted <MAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

20.6%; Score 115; DB 2; Length 514;

Best Local Similarity 95.2%; Pred. No. 2,55-05;

Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 52 WKNNRIWLOFAKLTGFTLMGR 72

DB 133 WKNNRIWLOFAKLTGFTLMGR 153

RESULT 3

UC7100

polygalacturonase Cha o 2 - Japanese cypress

C:Species: Chamaecyparis obtusa (Japanese cypress)

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-May-2000

C:Accession: JC7100; PC7026

R:Morit, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.

Biochem. Biophys. Res. Commun. 263, 166-171, 1999

A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all

A:Reference number: JC7100; MUID:99417540; PMID:10486272

A:Accession: JC7100

A:Molecule type: mRNA

A:Residues: 1-514 <MOR>

A:Accession: PC7026

A:Molecule type: protein

A:Residues: 51-62 <MO2>

Query Match

17.6%; Score 98; DB 2; Length 514;

Best Local Similarity 38.5%; Pred. No. 0.0023;

Matches 25; Conservative 8; Mismatches 16; Indels 16; Gaps 3;

QY 13 NRRVPIKRVSNVTHGRIDIFASKNFHLQKNTTGT-----GRRWKNNRIWLOFAKLTG 66

DB 98 NKKFV---VNNLVFRG-----PCQPHLSFKVDGTIVADPAPWAKSKITWLOFAQLTID 147

QY 67 FTLNG 71

DB 148 FTLNG 152

RESULT 4

JC2124

major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000

C:Accession: JC2124

R:Some, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A:Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A:Reference number: JC2123; MUID:94183234; PMID:8135802

A:Accession: JC2124

A:Molecule type: mRNA

A:Residues: 1-374 <SON>

A:Cross-references: GB:D26545; NID:G493633; PIDN:BA05543.1; PID:G493634

A:Experimental source: pollen

A:Note: the authors described carbohydrate binding site for residue 279

C:Superfamily: pectate lyase LATS9

C:Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>

F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

16.8%; Score 93.5; DB 2; Length 374;

Best Local Similarity 51.4%; Pred. No. 0.0054;

Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89

DB 133 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 153

RESULT 5

JC2123

major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000

C:Accession: JC2123; PC2065

R:Some, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A:Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

DB 85 RDRPLMIIFS-----GNNNIKLMKMPYIAGYKTFDGR 116

RESULT 5

JC2123

major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)

C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000

C:Accession: JC2123; PC2065

R:Some, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994

A:Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese

A:Reference number: JC2123; MUID:94183234; PMID:8135802

A:Accession: JC2123

A:Molecule type: mRNA

A:Residues: 1-374 <SON>

A:Cross-references: GB:D26544; NID:G493631; PIDN:BA05542.1; PID:G493632

A:Experimental source: pollen

A:Accession: PC2065

A:Molecule type: protein

A:Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <SO2>

A:Note: the authors described carbohydrate binding site for residue 279

C:Superfamily: pectate lyase LATS9

C:Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-374/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>

F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

16.8%; Score 93.5; DB 2; Length 374;

Best Local Similarity 51.4%; Pred. No. 0.0054;

Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGR 89

DB 85 RDRPLMIIFS-----GNNNIKLMKMPYIAGYKTFDGR 116

RESULT 6

JC7366

Jun a 2 protein - mountain cedar

C:Species: Juniperus ashei (mountain cedar)

C>Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 17-Nov-2000

C:Accession: JC7366; PC7093

R:Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.

Biochem. Biophys. Res. Commun. 275, 195-202, 2000

A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major alle

A:Reference number: JC7366

A:Accession: JC7366

A:Molecule type: mRNA

A:Residues: 1-507 <YOK>

A:Cross-references: GB:AJ404653

A:Accession: PC7093

A:Molecule type: protein

A:Residues: 55-63 <YO2>

C:Comment: This protein, a second major allergen of mountain cedar pollen, which is invo

o the polygalacturonase family.

C:Keywords: glycoprotein; pollen

Query Match

16.4%; Score 91.5; DB 2; Length 507;

Best Local Similarity 31.5%; Pred. No. 0.013;

Matches 29; Conservative 14; Mismatches 26; Indels 23; Gaps 5;

QY 12 PNRVPIKRVSNVTHGRIDIFASKNFHLQKNTTGT---GRRWKNNRIWLOFAKLTGFT 68

DB 97 PANKFV---VNNLVFRG---PCQPHLSFKVDGTIVADPAPWAKSKITWLOFAQLTID 150

QY 69 LMGRLKMPYIAGY-----KTFDGR 91

DB 151 LMGRLKMPYIAGY-----KTFDGR 177

RESULT 7

F72335

F72335

F72335

F72335

F72335

F72335

F72335

hypothetical protein TM0760 - Thermotoga maritima (strain MSB8)

C/Species: Thermotoga maritima

C/Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C/Accession: F72335

R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.

Nature 399, 323-329, 1999

A/Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A/Reference number: A72200; MUID:99287316; PMID:10360571

A/Accession: F72335

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-398 <ARN>

A/Cross-references: GB:AE001746; GB:AE000512; NID:g4981285; PIDN:AA035842.1; PID:g498128

A/Experimental source: strain MSB8

C/Genetics:

A/Genes: TM0760

Query Match

Best Local Similarity 15.1%; Score 8445; DB 2; Length 398;

Matches 25; Conservative 11; Mismatches 35; Indels 17; Gaps 2;

Db

28 GRRIDFASKNFHLOKNTIGTRKRNRIW-----LQPAKLTGFTLMGRRLMK 76

33 GHRVDIYIDFSLH-----TGRWSETFGWSFKDGVFIVERETRYGNSLSLSI 86

77 PMYIAGYTFDGRVDGIIAAYGNPASM 104

87 DYRNKRKTIQKRYVITANSRHPSP 114

RESULT 8

C84647

probable synaptochrome (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C/Accession: G94647

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.;

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: C84647

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-219 <STO>

A/Cross-references: GB:AE002093; NID:g4567243; PIDN:AA023657.1; GSPDB:GN00139

C/Genetics:

A/Genes: At2g25340

A/Map position: 2

Query Match

Best Local Similarity 14.5%; Score 81; DB 2; Length 219;

Matches 16; Conservative 14; Mismatches 22; Indels 2; Gaps 1;

Db

14 RRVFKRVSVIIGRRIDFASKNFHLOKNTIGTRKRNRIW-----LQPAKLTGFTLMGRRLMK 65

138 RQVMEINIDVLRGERELLDVKTANNQGTFRKQTRFNNTVMRNCKLT 191

RESULT 9

T04630

synaptochrome homolog F10N7.40 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-May-1999

C/Accession: T04630

R/Beyan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hohnselt, J.; Mewes, H.W.; Mayer, K.F.X

A/Residues: 1-219 <BRV>

A/Cross-references: EMBL:AL021636

A/Experimental source: cultivar Columbia; BAC clone F10N7

C/Genetics:

A/Map position: 4

A/Introns: 73/1; 136/3; 191/2

A/Note: F10N7.40

Query Match

Best Local Similarity 13.8%; Score 77; DB 2; Length 219;

Matches 22; Conservative 16; Mismatches 27; Indels 6; Gaps 3;

Db

14 RRVFKRVSVIIGRRIDFASKNFHLOKNTIGTRKRNRIW-----LQPAKLTGFTLMGRRLMK 70

138 RQVMEINIDVLRGERELLDVKTANNQGTFRKQTRFNNTVMRNCKLT--VILL 194

71 GRRLKMPVIA 81

195 ILLLVITIA 205

RESULT 10

S76496

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C/Species: Synechocystis sp.

A/Variety: PCC 6803 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C/Accession: S76496

R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

A/Reference number: S74322; MUID:97061201; PMID:8905231

A/Accession: S76496

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-477 <KAN>

A/Cross-references: EMBL:D90915; GB:AB001339; NID:g1653604; PIDN:BA18625.1; PID:g165371;

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Superfamily: starch synthase

Query Match

Best Local Similarity 13.5%; Score 75.5; DB 2; Length 477;

Matches 23; Conservative 15; Mismatches 29; Indels 19; Gaps 4;

Db

8 NQFQNRVYFKRVSVIIGRRIDFASKNFHLOKNTIGTRKRNRIW-----LQPAKLTGFTLMGRRLMK 67

31 HQGHVDFVFMPPYGF-----GDKIDV-----EKRPV-----WKGRMFOQFAVQSY 74

68 TLMGRRLKMPVITAGYKTFDGRVDG 93

75 L--PDTKIPLYLFGHPADSRRIYG 97

RESULT 11

AD3471

adenylate kinase (EC 2.7.1.25) (imported) - Brucella melitensis (strain 16M)

C/Species: Brucella melitensis

C/Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 19-Apr-2002

C/Accession: AD3471

R/Delvecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, I.

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A/Reference number: AD3252; PMID:11756688

A/Accession: AD3471

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-644 <KUD>

A/Cross-references: GB:AE009917; PIDN:AA52935.1; PID:g17983784; GSPDB:GN00190

A/Experimental source: strain 16M

C/Genetics:

A:Gene: BME11754  
A:Map position: 1  
C:Superfamily: nucleation protein nodQ; adenyllysulfate kinase homology; translation elc  
C:Keywords: phosphotransferase

Query Match 13.2%; Score 73.5; DB 2; Length 644;  
Best Local Similarity 26.0%; Pred. No. 2.1;  
Matches 26; Conservative 17; Mismatches 44; Indels 13; Gaps 4;

5 VAFNPGFNR---VFIRVSNVTHGRRID-IFASKNFHLQ-----KNTIGTRMK 53  
DB 402 IAFDYGERRATGAVLIDRLTNATVGAGMIDALRQNTVHDLQAFDLNQAARAGKFGK 461  
QY 54 NNRWLOPAKLTGFTLMGRRLKMPYIAGYKTF--DGRV 91  
DB 462 PAWLFTGTSASGKSTIANRLRQRLHALGKHTYLLDGDNV 501

## RESULT 12

A45049  
4-alpha-glucanotransferase (EC 2.4.1.25) - potato  
N:Alternate names: amylomalase; D-enzyme; disproportionating enzyme  
C:Species: Solanum tuberosum (potato)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A45049  
R:Takaha, T.; Yanase, M.; Okada, S.; Smith, S.M.  
J. Biol. Chem. 268, 1391-1396, 1993  
A:Title: Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25) of potato.  
A:Reference number: A45049; MUID:93123262; PMID:7678257  
A:Accession: A45049  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-576 <TAK>  
A:Cross-references: GB:X68664; GB:S52648; NID:9296691; PIDN:CAA48630.1; PID:9296692  
A:Experimental source: cv. May Queen, tubers  
A:Note: sequence extracted from NCBI backbone (NCBI:122077, NCBI:122078)  
C:Superfamily: 4-alpha-glucanotransferase  
C:Keywords: chloroplast; glycosyltransferase; hexosyltransferase

Query Match 12.7%; Score 71; DB 1; Length 576;  
Best Local Similarity 35.1%; Pred. No. 3.6;  
Matches 20; Conservative 10; Mismatches 13; Indels 14; Gaps 4;

QY 31 IIFPSKPNHLQKNTIGTRMKNNRIMLOPAKLTGFTLMGRRLKMPYIAGYKTFD 87  
DB 262 IDIFIAQQLFQ-----RQMKVYR---DYARSKGISIMG---DMPTIV-GYHSAD 304

## RESULT 13

E97154  
phosphopentomutase [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C:Accession: E97154  
R:Noiling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: E97154  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-330 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK80024.1; PID:g15025052; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC2065  
C:Superfamily: phosphopentomutase

Query Match 12.3%; Score 68.5; DB 2; Length 390;  
Best Local Similarity 29.0%; Pred. No. 4.5;  
Matches 27; Conservative 12; Mismatches 45; Indels 9; Gaps 4;

QY 6 AFNPGFNRVFIK-RVSNVTHGRRIDIPASKN-----FHLQKNTIGTR--WKNNRI 57  
DB 213 ALDPFNKTLKEIKNNKSNWCVGKIEDIFNKKYTEAVHILKNNMDGVDKLEYKTDKK 272  
QY 58 WLOPAKLTGFTLM-GRRLKMPYIAGYKTFDGR 89  
DB 273 GLITNLVDFTMLYGRNDPKGYAKALEFDFGR 305

## RESULT 14

S46755  
hypothetical protein YHR154w - Yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C>Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 19-Apr-2002  
C:Accession: S46755  
R:Macri, C.  
Submitted to the EMBL Data Library, June 1994  
A:Description: The sequence of S. cerevisiae cosmid 9666.  
A:Reference number: S46752  
A:Accession: S46755  
A:Molecule type: DNA  
A:Residues: 1-1070 <MAC>  
A:Cross-references: EMBL:U10397; NID:9500647; PID:9500651; GSPDB:GN00008; MIPS:YHR154w  
C:Genetics:  
A:Gene: SCD; MIPS:YHR154w  
A:Cross-references: SCD:S0001197  
A:Map position: 8R

Query Match 12.1%; Score 67.5; DB 2; Length 1070;  
Best Local Similarity 27.4%; Pred. No. 18;  
Matches 20; Conservative 15; Mismatches 23; Indels 15; Gaps 4;

QY 2 KTVAFNPGFNRVFIKRVSNVTHGRRIDIPASKNHLQKNTIG-----TGRWK--- 53  
DB 379 ELTVAAVTNFGSGRFQRLV-ILGSLTPELTRKNTLITKSTIGKKFVAKKSLDP 437  
QY 54 -----NNRIMLO 60  
DB 438 QNAIIVTNHMLE 450

## RESULT 15

JM0107  
very-long-chain acyl-CoA synthetase related protein - mouse  
N:Alternate names: VLACSR  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 08-Sep-2000  
C:Accession: JM0107  
R:Berger, J.; Truppe, C.; Neumann, H.; Forss-Peterson, S.  
Biochem. Biophys. Res. Commun. 247, 255-260, 1998  
A:Title: A novel relative of the very-long-chain acyl-CoA synthetase and fatty acid trans  
A:Reference number: JM0107; MUID:98308102; PMID:9642112  
A:Accession: JM0107  
A:Molecule type: mRNA  
A:Residues: 1-689 <BER>  
A:Cross-references: GB:AJ223959  
A:Experimental source: liver  
C:Comment: This protein likely functions as a plasma membrane transporter of long chain f  
cids.  
C:Superfamily: Mycobacterium tuberculosis probable fad6 protein; acetate-CoA ligase hom  
F:169-647/Domain: acetate-CoA ligase homology <ACL>

Query Match 12.0%; Score 67; DB 1; Length 689;  
Best Local Similarity 21.7%; Pred. No. 13;  
Matches 25; Conservative 24; Mismatches 42; Indels 24; Gaps 4;

QY 8 NOFGFNRV--IKRVSNVTHGRRIDIPASKNHLQKNTIGTRRRKNNRI----- 57  
DB 510 SQASNRKLVANRRVGDLYFNNGDVLTDQEGFFYFQRLDTPFKNGENSTGEVSV 569  
QY 58 -----WLOPAKLTGFTLMGRRLKMPYIAGYKTFDGRVDGIIAAYONPASW 104  
DB 570 LSSIDFLFEEVNVGVVPVCGCGKAVMAAVLAPGKTFDQKL-----YOHVRSW 618

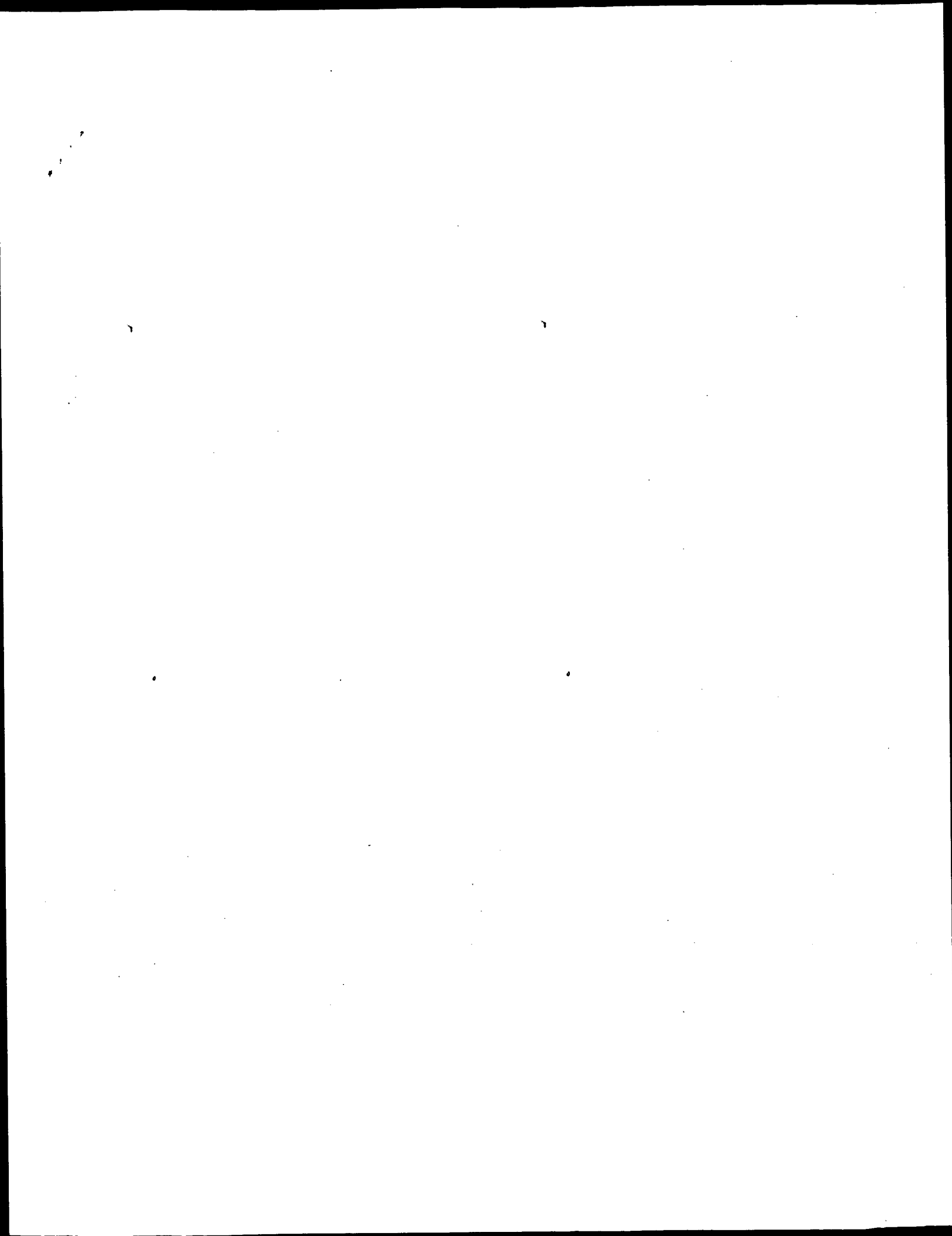
Wed Apr 23 18:08:51 2003

us-09-142-524d-2.rpt

Page 5

Search completed: April 20, 2003, 13:15:37  
Job time : 45.5526 secs

---





GenCore version 5.1.4 ps 4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 20.4474 Seconds

(without alignments)  
212.987 Million cell updates/sec

Title: US-09-142-524D-2

Sequence: 1 MKVTVAENQFGFNREVFTR.....FDGRRVDGIIAAYQNPASWK 105

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115.5	20.7	514	1	MPA2_CRYVA
2	93.5	16.8	374	1	SBP_CRYVA
3	85.5	15.3	367	1	MPA1_TJNNA
4	82.5	14.8	346	1	MPA1_CUPAN
5	75.5	13.5	477	1	GLGA_SYNY3
6	74.5	13.4	375	1	MPA1_CHAOB
7	71	12.7	576	1	DEEP_SOLTU
8	68.5	12.3	234	1	A29B_DROME
9	67.5	12.1	1070	1	YHV4_YEAST
10	66.5	11.9	110	1	Y12K_YEAST
11	66	11.8	196	1	YMO7_YEAST
12	65.5	11.7	493	1	AMVR_DROSE
13	65.5	11.7	699	1	NUSC_DIGGR
14	64.5	11.6	493	1	AMVR_DROOR
15	63.5	11.4	130	1	RS9_NEIMA
16	63.5	11.4	542	1	LAC3_THACU
17	63	11.3	570	1	Z136_HUMAN
18	63	11.3	880	1	YMD5_CABEL
19	62.5	11.2	449	1	PE59_LYCES
20	62.5	11.2	493	1	AMVR_DROBE
21	62.5	11.2	493	1	AMVR_DROME
22	62.5	11.2	493	1	AMVR_DROTE
23	62.5	11.2	493	1	AMVR_DROTE
24	62.5	11.2	493	1	AMVR_DROTE
25	62	11.1	1121	1	MDR6_DROYA
26	61.5	11.0	493	1	AMVR_DROER
27	61.5	11.0	646	1	NODO_RHIS3
28	61.5	11.0	781	1	TL22_CHICK
29	61.5	11.0	2258	1	PAS_FIG
30	61	10.9	275	1	VNS2_DSDNV
31	61	10.9	505	1	CSA5_DROME
32	60.5	10.8	633	1	NODO_RHISB
33	60.5	10.8	706	1	KPCT_HUMAN

34	60.5	10.8	1010	1	CLPP_CHLEU
35	60.5	10.8	2133	1	PAB_FIG
36	60	10.8	130	1	RS9_XYLEFA
37	60	10.8	262	1	FLGG_AGRFS
38	60	10.8	275	1	VNS2_JCDNV
39	59.5	10.7	549	1	LIP2_CANRU
40	59.5	10.7	718	1	CDG1_BACSS
41	59.5	10.7	826	1	YEHF_ECOLI
42	59.5	10.6	461	1	TBG_NEUCR
43	59	10.6	740	1	GNT5_RAT
44	59	10.6	741	1	GNT5_RAT
45	59	10.6	842	1	PHS2_RAT

## ALIGNMENTS

RESULT 1  
MPA2\_CRYVA STANDARD; PRT; 514 AA.  
AC P43212;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Possible polygalacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)  
DE (Major pollen allergen Cry j 2) (Cry j II)  
OS Cryptomeria japonica (Japanese cedar).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferales; Cupressaceae; Cryptomeria.  
OX NCBI\_TaxID=3369;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Pollen;  
RX MEDLINE=95010777; PubMed=7926035;  
RA Namba M., Kurose M., Toriyo K., Hino K., Taniguchi Y., Fukuda S.,  
RA Usui M., Kurimoto M.;  
RT "Molecular cloning of the second major allergen, Cry j II, from  
RT Japanese cedar pollen.";  
RL PNAS Lett. 353:124-128 (1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pollen;  
RX MEDLINE=94271186; PubMed=8002972;  
RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;  
RT "cDNA cloning and expression of Cry j II the second major allergen of  
RT Japanese cedar pollen.";  
RL Biochem. Biophys. Res. Commun. 201:1021-1028 (1994).  
RN [3]  
RP SEQUENCE OF 55-64.  
RX MEDLINE=90342988; PubMed=2382797;  
RA Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhasi T.;  
RT "Identification of the second major allergen of Japanese cedar  
RT pollen.";  
RL Allergy 45:309-312 (1990).  
RN [4]  
RP CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-  
RP galactosiduronic linkages in pectate and other galacturonans.  
RP -1- SUPERFICIAL LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).  
RP -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES  
RP (POLYGALACTURONASES).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>)  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk).  
DR EMBL, D3765; BA07021.1; -;  
DR EMBL, D29772; BA06172.1; -;  
DR HSRP, P26509; IBBE.  
DR InterPro, IPR00743; GR28.  
DR Pfam, PF00295; Glyco\_hydro\_28; 1.

RX	* HJLJKETA A., Matsunoto I., /Ozuna K., Ogawa H.; Medline=95003748; PubMed=7608114;
RA	"Antigenicity of the oligosaccharide moiety of the Japanese cedar (Cryptomeria japonica) pollen allergen, Cry jI."; <i>Int. Arch. Allergy Immunol.</i> 105:198-202(1994).
RL	[5]
RP	STRUCTURE OF CARBOHYDRATES.
RC	Tissue=Pollen;
RX	MEDLINE=95332249; Pubmed=7608114;
RA	Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M., Fukuda S., Hanazawa H., Haruyama H., Kurimoto M. "Carbohydrate structures of the glycoprotein allergen Cry j I from Japanese cedar ( <i>Cryptomeria japonica</i> ) pollen." <i>J. Biochem.</i> 117:289-295(1995).
RT	-1 - PM: CONTAINS FUCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC	-1 DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR POLLEN.
CC	POLLEN, THE MOST COMMON POLEEN ALLERGEN IN JAPAN.
CC	-1 MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM B DIFFERS IN SIX POSITIONS.
CC	-1 SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC	Amb A I/Amb A II/Cry J I SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch). -----
DR	EMBL; D26544; BAA05542.1; .-
DR	EMBL; D26545; BAA05543.1; .-
DR	EMBL; D34639; BAA07020.1; .-
DR	PIR; A44773; A44773.
DR	GlycoSuiteDB; P18632; -. Amb_allergen.
DR	IInterPro; IPRO02022; Amb_1.
DR	pfam; PF00544; pec_lyase; 1.
DR	PRINTS; PR00807; AMBALLERGEN.
KW	Allergen; Glycoprotein; Multigene family; Signal.
FT	SIGNAL 1 21
FT	CHAIN 22 374
FT	CARBOHD 158 158      SUGI BASIC PROTEIN.    ) (POTENTIAL)
FT	CARBOHD 191 191      N-LINKED (GLCNAG.. ) (COMPLEX).
FT	/FTID=CAR 000135.
FT	N-LINKED (GLCNAc.. ) (POTENTIAL)
FT	CARBOHD 293 293      N-LINKED (GLCNAc.. ) (COMPLEX).
FT	CARBOHD 354 354      FTID=CAR 000136.
FT	VARIANT 12 12      L-> F (IN CRy J 1-B).
FT	VARIANT 143 143      H-> V (IN CRy J 1-B).
FT	VARIANT 202 202      S-> T (IN CRy J 1-B).
FT	VARIANT 221 221      L-> S (IN CRy J 1-B).
FT	VARIANT 358 358      Q-> H (IN CRy J 1-B).
FT	VARIANT 361 361      K-> Q (IN CRy J 1-B).
SQ	SEQUENCE 374 AA, 40645 MW, 74ABZ950248F56F CRC64;
DY	Query Match 16.8%; Score 93.5; DB 1; Length 374; Best Local Similarity 51.4%; Pred. No. 0.0022; Matches 19; Conservative 6; Mismatches 7; Indels 5; Gaps 1
ID	MPAL JUNAS STANDARD; PRT; 367 AA.
AC	F81294; Q9ZNUT7.
DT	16-OCT-2001 (Rel. 40, Created) 16-OCT-2001 (Rel. 40, Last sequence update) 15-JUN-2002 (Rel. 41, Last annotation update) Major pollen allergen Jun a 1 precursor.
OS	Juniperus abelii (Ozak white cedar).

[illegible]

```

EMBL; D90915; BAA18625.1; -.
DR InterPro; IPR001296; Glycos. transf. 1.
DR Pfam; PF00534; Glycos. transf. 1; 1.
KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
KW Complete proteome.
FT BINDING 15 ADP-GLUCOSE (BY SIMILARITY).
SQ SEQUENCE 477 AA; 54346 MW; 40C1B6013938E032 CRC64;

Query Match 13.5%; Score 75.5; DB 1; Length 477;
Best Local Similarity 26.7%; Pred. No. 0.35;
Matches 23; Conservative 15; Mismatches 29; Indels 19; Gaps 4;

QY 8 NQGPENRVRVFIKRVSNVHIGRRIDIPASKNFHLQKNTGTGRRWKNRRWLPQAKITGF 67
DB 31 HOLGHADVRFVPPYGFIL---GDKTIDV-----PKRPV-----WKGRAMQGFVAVQSY 74
QY 68 TLNGRRLLKMPYIAGYKTFDGRVGD 93
DB 75 L---PDTKIDPLYLFGHPAFDSRRITYG 97

RESULT 6
MPAL_CHAOB STANDARD; PRT; 375 AA.
ID MPAL_CHAOB
AC Q96385;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major pollen allergen Cha o 1 precursor.
OS Chamaecyparis obtusa (Japanese cypress).
OC Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;
OC Chamaecyparis.
OC NCBI_Taxid=13415;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=96265194; Pubmed=8676896;
RA Suzuki M., Komiyaama N., Itoh M., Itoh H., Sone T., Kuno K., Takagi I.,
RA Ohta N.;
RT "Purification, characterization and molecular cloning of Cha o 1, a
RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";
RL Mol. Immunol. 33:451-460(1996).
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D45404; BAA08246.1; -.
DR InterPro; IPR002022; Amb allergen.
DR Pfam; PF00544; pec lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Allergen; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 22 MAJOR POLLEN ALLERGEN CHA O 1.
FT CARBOHYD 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF7065DBF CRC64;

Query Match 13.4%; Score 74.5; DB 1; Length 375;
Best Local Similarity 40.5%; Pred. No. 0.35;
Matches 15; Conservative 7; Mismatches 10; Indels 5; Gaps 1.

```

Db 85 RERSLWIFSK-----NLNKLNNPLVYAGNKTTIDGR 116

RESULT 7  
DPEP SOLUTU STANDARD; PRT; 576 AA.  
AC 006801;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 4-alpha-glucanotransferase, chloroplast precursor (EC 2.4.1.25)  
DN (Amylomylase) (Disproportionating enzyme) (D-enzyme).  
GN DPEP.  
OS Solanum tuberosum (Potato); Streptophyta; Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Spermatophyta; Magnoliophyta; eucolecledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxId=4113;  
[1]  
SEQUENCE FROM N.A., AND SEQUENCE OF 53-57; 174-183 AND 247-259.  
RP STRAIN=cv. May Queen; TISSUE=tuber;  
FX MEDLINE=93123262; PubMed=7678257;  
RA Takaha T., Yanase M., Okada S., Smith S.M.;  
RT "Disproportionating enzyme (4-alpha-glucanotransferase; EC 2.4.1.25)  
of potato. Purification, molecular cloning, and potential role in  
starch metabolism.";  
RL J. Biol. Chem. 268:1391-1396(1993)).  
CC - FUNCTION: MAY ACT DURING STARCH BREAKDOWN TO CONVERT SMALL  
OLIGOSACCHARIDES INTO LARGER MOLECULES UPON WHICH STARCH  
PHOSPHORYLASE CAN ACT. OR MAY CHANGE THE STRUCTURE OF STARCH  
MOLECULES AND GRAIN ARCHITECTURE BY MODIFYING CHAIN LENGTH. OR  
MAY GENERATE FROM STARCH AND GLUCOSE OLIGOSACCHARIDES WHICH  
CAN SERVE EITHER AS PRIMERS FOR NEW STARCH PHOSPHOENZYME.  
CC - CATALYTIC ACTIVITY: Transfers a segment of a (1,4)-alpha-D-glucan  
to a new 4-position in an acceptor, which may be glucose or (1,4)-  
alpha-D-glucan.  
CC - SUBCELLULAR LOCATION: CHLOROPLAST OR AMYLOPLAST.  
CC - TISSUE SPECIFICITY: PRESENT IN LEAVES, STEMS, ROOTS, AND STOLONS  
BUT IS MOST ABUNDANT IN DEVELOPING AND MATURE TUBERS.  
CC - SIMILARITY: BELONGS TO THE DISPROPORTIONATING ENZYME FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>)  
or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk).  
-----  
DR EMBL; X68664; GAA48630.1; --  
DR PIR; A45049; A45049.  
DR HSP; O87172; 1CMTY.  
DR InterPro; IPRO03585; 4A\_glucoamylase.  
DR Pfam; PF02446; 4A\_glucoamylase; 1.  
DR TIGRPFAMs; TIGR00217; malQ; 1.  
KW Transferase; Glycosyltransferase; Carbohydrate metabolism;  
KM Amyloplast; Chloroplast; Transit peptide.  
FT TRANSIT 1 52 CHLOROPLAST.  
FT CHAIN 53 576 4-ALPHA-GLUCANOTRANSFERASE.  
SQ SEQUENCE 576 AA; 64950 MW; ADD16F3A546307BB CRC64.

Query Match 12.7%; Score 71; DB 1; Length 576;  
Best local Similarity 35.1%; Pred. No. 1.4;  
Matches 20; Conservative 10; Mismatches 13; Indels 14; Gaps 4.

Oy 31 IDIFAKNHLDKNTGTGGRKKNRIMVDFAKLTGFTLMGRKMKPMVYAGKYFD 87  
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
Db 262 IDIFAQQLFQ-----RWKRVR--DYARKSGISIMG--DWPIIV-GHNSAD 304

RESULT 8  
A29B\_DROME

ID A298.DROME STANDARD; PRT; 234 AA.  
 AC Q6197; Q9VT3; Q9TW05; Q9TW06; Q9TW07; Q9U976; Q9U977; Q9U978;  
 AC Q9U979; Q9U905;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Accessory gland protein Acp29Ab precursor.  
 GN ACP29AB OR CG17797.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pelecyota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_Taxid=7227;  
 RX MEDLINE=1227;  
 RP [1]  
 RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
 RC STRAIN=canon-S; TISSUE=Male accessory gland;  
 RX MEDLINE=98135120; PubMed=9474779;  
 RA Wolfer M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,  
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Tzeng U.K.;  
 RT "New genes for male accessory gland proteins in Drosophila  
 RT melanogaster";  
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Various strains;  
 RX MEDLINE=9282496; PubMed=10353698;  
 RA Agnede M.;  
 RT "Positive selection drives the evolution of the Acp29Ab accessory  
 RT gland protein in Drosophila";  
 RL Genetics 152:543-551(1999).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Various strains;  
 RX MEDLINE=20556153; PubMed=1102381;  
 RA Begun D.J., Whitley P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;  
 RT "Molecular population genetics of male accessory gland proteins in  
 RT Drosophila";  
 RL Genetics 156:1879-1888(2000).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Gelinkher S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer G.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Bensu P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Borkov A.D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Cantler A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.B., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz N.S., Gelbart W.M., Glasser K.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gerhart M., Fleischmann W.,  
 RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao X., Li X., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Lin X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclib J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R., Lillard J., Puri V., Reese M.G.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtk R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
 CC MATED FEMALE FLIES.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLAND AND IN  
 CC SEMINAL FLUID.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U85758; AAB96382.1; -  
 DR EMBL; AJ240513; CAB53187.1; -  
 DR EMBL; AJ240514; CAB53188.1; -  
 DR EMBL; AJ240515; CAB53189.1; -  
 DR EMBL; AJ240516; CAB53190.1; -  
 DR EMBL; AJ240517; CAB53191.1; -  
 DR EMBL; AJ240518; CAB53192.1; -  
 DR EMBL; AJ240519; CAB53193.1; -  
 DR EMBL; AJ240520; CAB53194.1; -  
 DR EMBL; AJ240521; CAB53195.1; -  
 DR EMBL; AJ240522; CAB53196.1; -  
 DR EMBL; AJ240523; CAB53197.1; -  
 DR EMBL; AJ240524; CAB53198.1; -  
 DR EMBL; AJ240525; CAB53199.1; -  
 DR EMBL; AJ240526; CAB53200.1; -  
 DR EMBL; AJ240527; CAB53201.1; -  
 DR EMBL; AJ240528; CAB53202.1; -  
 DR EMBL; AJ240529; CAB53203.1; -  
 DR EMBL; AJ240530; CAB53204.1; -  
 DR EMBL; AJ240531; CAB53205.1; -  
 DR EMBL; AJ240532; CAB53206.1; -  
 DR EMBL; AJ240533; CAB53207.1; -  
 DR EMBL; AJ240534; CAB53208.1; -  
 DR EMBL; AJ240535; CAB53209.1; -  
 DR EMBL; AJ240536; CAB53210.1; -  
 DR EMBL; AJ240537; CAB53211.1; -  
 DR EMBL; AJ240538; CAB53212.1; -  
 DR EMBL; AJ240539; CAB53213.1; -  
 DR EMBL; AJ240540; CAB53214.1; -  
 DR EMBL; AJ240541; CAB53215.1; -  
 DR EMBL; AJ240542; CAB53216.1; -  
 DR EMBL; AJ240543; CAB53217.1; -  
 DR EMBL; AJ240544; CAB53218.1; -  
 DR EMBL; AJ240545; CAB53219.1; -  
 DR EMBL; AJ240546; CAB53220.1; -  
 DR EMBL; AJ240547; CAB53221.1; -  
 DR EMBL; AJ240548; CAB53222.1; -  
 DR EMBL; AJ240549; CAB53223.1; -  
 DR EMBL; AJ240550; CAB53224.1; -  
 DR EMBL; AJ240551; CAB53225.1; -  
 DR EMBL; AJ240552; CAB53226.1; -  
 DR EMBL; AJ240553; CAB53227.1; -  
 DR EMBL; AJ240554; CAB53228.1; -  
 DR EMBL; AJ240555; CAB53229.1; -  
 DR EMBL; AJ240556; CAB53230.1; -  
 DR EMBL; AJ240557; CAB53231.1; -  
 DR EMBL; AJ240558; CAB53232.1; -  
 DR EMBL; AJ240559; CAB53233.1; -  
 DR EMBL; AJ240560; CAB53234.1; -  
 DR EMBL; AJ240561; CAB53235.1; -  
 DR EMBL; AJ240562; CAB53236.1; -  
 DR EMBL; AJ240563; CAB53237.1; -  
 DR EMBL; AJ240564; CAB53238.1; -  
 DR EMBL; AJ240565; CAB53239.1; -  
 DR EMBL; AJ240566; CAB53240.1; -  
 DR EMBL; AJ240567; CAB53241.1; -  
 DR EMBL; AJ240568; CAB53242.1; -  
 DR EMBL; AJ240569; CAB53243.1; -  
 DR EMBL; AJ240570; CAB53244.1; -  
 DR EMBL; AJ240571; CAB53245.1; -  
 DR EMBL; AJ240572; CAB53246.1; -  
 DR EMBL; AJ240573; CAB53247.1; -  
 DR EMBL; AJ240574; CAB53248.1; -  
 DR EMBL; AJ240575; CAB53249.1; -  
 DR EMBL; AJ240576; CAB53250.1; -  
 DR EMBL; AJ240577; CAB53251.1; -  
 DR EMBL; AJ240578; CAB53252.1; -  
 DR EMBL; AJ240579; CAB53253.1; -  
 DR EMBL; AJ240580; CAB53254.1; -  
 DR EMBL; AJ240581; CAB53255.1; -  
 DR EMBL; AJ240582; CAB53256.1; -  
 DR EMBL; AJ240583; CAB53257.1; -  
 DR EMBL; AJ240584; CAB53258.1; -  
 DR EMBL; AJ240585; CAB53259.1; -  
 DR EMBL; AJ240586; CAB53260.1; -  
 DR EMBL; AJ240587; CAB53261.1; -  
 DR EMBL; AJ240588; CAB53262.1; -  
 DR EMBL; AJ240589; CAB53263.1; -  
 DR EMBL; AJ240590; CAB53264.1; -  
 DR EMBL; AJ240591; CAB53265.1; -  
 DR EMBL; AJ240592; CAB53266.1; -  
 DR EMBL; AJ240593; CAB53267.1; -  
 DR EMBL; AJ240594; CAB53268.1; -  
 DR EMBL; AJ240595; CAB53269.1; -  
 DR EMBL; AJ240596; CAB53270.1; -  
 DR EMBL; AJ240597; CAB53271.1; -  
 DR EMBL; AJ240598; CAB53272.1; -  
 DR EMBL; AJ240599; CAB53273.1; -  
 DR EMBL; AJ240600; CAB53274.1; -  
 DR EMBL; AJ240601; CAB53275.1; -  
 DR EMBL; AJ240602; CAB53276.1; -  
 DR EMBL; AJ240603; CAB53277.1; -  
 DR EMBL; AJ240604; CAB53278.1; -  
 DR EMBL; AJ240605; CAB53279.1; -  
 DR EMBL; AJ240606; CAB53280.1; -  
 DR EMBL; AJ240607; CAB53281.1; -  
 DR EMBL; AJ240608; CAB53282.1; -  
 DR EMBL; AJ240609; CAB53283.1; -  
 DR EMBL; AJ240610; CAB53284.1; -  
 DR EMBL; AJ240611; CAB53285.1; -  
 DR EMBL; AJ240612; CAB53286.1; -  
 DR EMBL; AJ240613; CAB53287.1; -  
 DR EMBL; AJ240614; CAB53288.1; -  
 DR EMBL; AJ240615; CAB53289.1; -  
 DR EMBL; AJ240616; CAB53290.1; -  
 DR EMBL; AJ240617; CAB53291.1; -  
 DR EMBL; AJ240618; CAB53292.1; -  
 DR EMBL; AJ240619; CAB53293.1; -  
 DR EMBL; AJ240620; CAB53294.1; -  
 DR EMBL; AJ240621; CAB53295.1; -  
 DR EMBL; AJ240622; CAB53296.1; -  
 DR EMBL; AJ240623; CAB53297.1; -  
 DR EMBL; AJ240624; CAB53298.1; -  
 DR EMBL; AJ240625; CAB53299.1; -  
 DR EMBL; AJ240626; CAB53300.1; -  
 DR EMBL; AJ240627; CAB53301.1; -  
 DR EMBL; AJ240628; CAB53302.1; -  
 DR EMBL; AJ240629; CAB53303.1; -  
 DR EMBL; AJ240630; CAB53304.1; -  
 DR EMBL; AJ240631; CAB53305.1; -  
 DR EMBL; AJ240632; CAB53306.1; -  
 DR EMBL; AJ240633; CAB53307.1; -  
 DR EMBL; AJ240634; CAB53308.1; -  
 DR EMBL; AJ240635; CAB53309.1; -  
 DR EMBL; AJ240636; CAB53310.1; -  
 DR EMBL; AJ240637; CAB53311.1; -  
 DR EMBL; AJ240638; CAB53312.1; -  
 DR EMBL; AJ240639; CAB53313.1; -  
 DR EMBL; AJ240640; CAB53314.1; -  
 DR EMBL; AJ240641; CAB53315.1; -  
 DR EMBL; AJ240642; CAB53316.1; -  
 DR EMBL; AJ240643; CAB53317.1; -  
 DR EMBL; AJ240644; CAB53318.1; -  
 DR EMBL; AJ240645; CAB53319.1; -  
 DR EMBL; AJ240646; CAB53320.1; -  
 DR EMBL; AJ240647; CAB53321.1; -  
 DR EMBL; AJ240648; CAB53322.1; -  
 DR EMBL; AJ240649; CAB53323.1; -  
 DR EMBL; AJ240650; CAB53324.1; -  
 DR EMBL; AJ240651; CAB53325.1; -  
 DR EMBL; AJ240652; CAB53326.1; -  
 DR EMBL; AJ240653; CAB53327.1; -  
 DR EMBL; AJ240654; CAB53328.1; -  
 DR EMBL; AJ240655; CAB53329.1; -  
 DR EMBL; AJ240656; CAB53330.1; -  
 DR EMBL; AJ240657; CAB53331.1; -  
 DR EMBL; AJ240658; CAB53332.1; -  
 DR EMBL; AJ240659; CAB53333.1; -  
 DR EMBL; AJ240660; CAB53334.1; -  
 DR EMBL; AJ240661; CAB53335.1; -  
 DR EMBL; AJ240662; CAB53336.1; -  
 DR EMBL; AJ240663; CAB53337.1; -  
 DR EMBL; AJ240664; CAB53338.1; -  
 DR EMBL; AJ240665; CAB53339.1; -  
 DR EMBL; AJ240666; CAB53340.1; -  
 DR EMBL; AJ240667; CAB53341.1; -  
 DR EMBL; AJ240668; CAB53342.1; -  
 DR EMBL; AJ240669; CAB53343.1; -  
 DR EMBL; AJ240670; CAB53344.1; -  
 DR EMBL; AJ240671; CAB53345.1; -  
 DR EMBL; AJ240672; CAB53346.1; -  
 DR EMBL; AJ240673; CAB53347.1; -  
 DR EMBL; AJ240674; CAB53348.1; -  
 DR EMBL; AJ240675; CAB53349.1; -  
 DR EMBL; AJ240676; CAB53350.1; -  
 DR EMBL; AJ240677; CAB53351.1; -  
 DR EMBL; AJ240678; CAB53352.1; -  
 DR EMBL; AJ240679; CAB53353.1; -  
 DR EMBL; AJ240680; CAB53354.1; -  
 DR EMBL; AJ240681; CAB53355.1; -  
 DR EMBL; AJ240682; CAB53356.1; -  
 DR EMBL; AJ240683; CAB53357.1; -  
 DR EMBL; AJ240684; CAB53358.1; -  
 DR EMBL; AJ240685; CAB53359.1; -  
 DR EMBL; AJ240686; CAB53360.1; -  
 DR EMBL; AJ240687; CAB53361.1; -  
 DR EMBL; AJ240688; CAB53362.1; -  
 DR EMBL; AJ240689; CAB53363.1; -  
 DR EMBL; AJ240690; CAB53364.1; -  
 DR EMBL; AJ240691; CAB53365.1; -  
 DR EMBL; AJ240692; CAB53366.1; -  
 DR EMBL; AJ240693; CAB53367.1; -  
 DR EMBL; AJ240694; CAB53368.1; -  
 DR EMBL; AJ240695; CAB53369.1; -  
 DR EMBL; AJ240696; CAB53370.1; -  
 DR EMBL; AJ240697; CAB53371.1; -  
 DR EMBL; AJ240698; CAB53372.1; -  
 DR EMBL; AJ240699; CAB53373.1; -  
 DR EMBL; AJ240700; CAB53374.1; -  
 DR EMBL; AJ240701; CAB53375.1; -  
 DR EMBL; AJ240702; CAB53376.1; -  
 DR EMBL; AJ240703; CAB53377.1; -  
 DR EMBL; AJ240704; CAB53378.1; -  
 DR EMBL; AJ240705; CAB53379.1; -  
 DR EMBL; AJ240706; CAB53380.1; -  
 DR EMBL; AJ240707; CAB53381.1; -  
 DR EMBL; AJ240708; CAB53382.1; -  
 DR EMBL; AJ240709; CAB53383.1; -  
 DR EMBL; AJ240710; CAB53384.1; -  
 DR EMBL; AJ240711; CAB53385.1; -  
 DR EMBL; AJ240712; CAB53386.1; -  
 DR EMBL; AJ240713; CAB53387.1; -  
 DR EMBL; AJ240714; CAB53388.1; -  
 DR EMBL; AJ240715; CAB53389.1; -  
 DR EMBL; AJ240716; CAB53390.1; -  
 DR EMBL; AJ240717; CAB53391.1; -  
 DR EMBL; AJ240718; CAB53392.1; -  
 DR EMBL; AJ240719; CAB53393.1; -  
 DR EMBL; AJ240720; CAB53394.1; -  
 DR EMBL; AJ240721; CAB53395.1; -  
 DR EMBL; AJ240722; CAB53396.1; -  
 DR EMBL; AJ240723; CAB53397.1; -  
 DR EMBL; AJ240724; CAB53398.1; -  
 DR EMBL; AJ240725; CAB53399.1; -  
 DR EMBL; AJ240726; CAB53400.1; -  
 DR EMBL; AJ240727; CAB53401.1; -  
 DR EMBL; AJ240728; CAB53402.1; -  
 DR EMBL; AJ240729; CAB53403.1; -  
 DR EMBL; AJ240730; CAB53404.1; -  
 DR EMBL; AJ240731; CAB53405.1; -  
 DR EMBL; AJ240732; CAB53406.1; -  
 DR EMBL; AJ240733; CAB53407.1; -  
 DR EMBL; AJ240734; CAB53408.1; -  
 DR EMBL; AJ240735; CAB53409.1; -  
 DR EMBL; AJ240736; CAB53410.1; -  
 DR EMBL; AJ240737; CAB53411.1; -  
 DR EMBL; AJ240738; CAB53412.1; -  
 DR EMBL; AJ240739; CAB53413.1; -  
 DR EMBL; AJ240740; CAB53414.1; -  
 DR EMBL; AJ240741; CAB53415.1; -  
 DR EMBL; AJ240742; CAB53416.1; -  
 DR EMBL; AJ240743; CAB53417.1; -  
 DR EMBL; AJ240744; CAB53418.1; -  
 DR EMBL; AJ240745; CAB53419.1; -  
 DR EMBL; AJ240746; CAB53420.1; -  
 DR EMBL; AJ240747; CAB53421.1; -  
 DR EMBL; AJ240748; CAB53422.1; -  
 DR EMBL; AJ240749; CAB53423.1; -  
 DR EMBL; AJ240750; CAB53424.1; -  
 DR EMBL; AJ240751; CAB53425.1; -  
 DR EMBL; AJ240752; CAB53426.1; -  
 DR EMBL; AJ240753; CAB53427.1; -  
 DR EMBL; AJ240754; CAB53428.1; -  
 DR EMBL; AJ240755; CAB53429.1; -  
 DR EMBL; AJ240756; CAB53430.1; -  
 DR EMBL; AJ240757; CAB53431.1; -  
 DR EMBL; AJ240758; CAB53432.1; -  
 DR EMBL; AJ240759; CAB53433.1; -  
 DR EMBL; AJ240760; CAB53434.1; -  
 DR EMBL; AJ240761; CAB53435.1; -  
 DR EMBL; AJ240762; CAB53436.1; -  
 DR EMBL; AJ240763; CAB53437.1; -  
 DR EMBL; AJ240764; CAB53438.1; -  
 DR EMBL; AJ240765; CAB53439.1; -  
 DR EMBL; AJ240766; CAB53440.1; -  
 DR EMBL; AJ240767; CAB53441.1; -  
 DR EMBL; AJ240768; CAB53442.1; -  
 DR EMBL; AJ240769; CAB53443.1; -  
 DR EMBL; AJ240770; CAB53444.1; -  
 DR EMBL; AJ240771; CAB53445.1; -  
 DR EMBL; AJ240772; CAB53446.1; -  
 DR EMBL; AJ240773; CAB53447.1; -  
 DR EMBL; AJ240774; CAB53448.1; -  
 DR EMBL; AJ240775; CAB53449.1; -  
 DR EMBL; AJ240776; CAB53450.1; -  
 DR EMBL; AJ240777; CAB53451.1; -  
 DR EMBL; AJ240778; CAB53452.1; -  
 DR EMBL; AJ240779; CAB53453.1; -  
 DR EMBL; AJ240780; CAB53454.1; -  
 DR EMBL; AJ240781; CAB53455.1; -  
 DR EMBL; AJ240782; CAB53456.1; -  
 DR EMBL; AJ240783; CAB53457.1; -  
 DR EMBL; AJ240784; CAB53458.1; -  
 DR EMBL; AJ240785; CAB53459.1; -  
 DR EMBL; AJ240786; CAB53460.1; -  
 DR EMBL; AJ240787; CAB53461.1; -  
 DR EMBL; AJ240788; CAB53462.1; -  
 DR EMBL; AJ240789; CAB53463.1; -  
 DR EMBL; AJ240790; CAB53464.1; -  
 DR EMBL; AJ240791; CAB53465.1; -  
 DR EMBL; AJ240792; CAB53466.1; -  
 DR EMBL; AJ240793; CAB53467.1; -  
 DR EMBL; AJ240794; CAB53468.1; -  
 DR EMBL; AJ240795; CAB53469.1; -  
 DR EMBL; AJ240796; CAB53470.1; -  
 DR EMBL; AJ240797; CAB53471.1; -  
 DR EMBL; AJ240798; CAB53472.1; -  
 DR EMBL; AJ240799; CAB53473.1; -  
 DR EMBL; AJ240800; CAB53474.1; -  
 DR EMBL; AJ240801; CAB53475.1; -  
 DR EMBL; AJ240802; CAB53476.1; -  
 DR EMBL; AJ240803; CAB53477.1; -  
 DR EMBL; AJ240804; CAB53478.1; -  
 DR EMBL; AJ240805; CAB53479.1; -  
 DR EMBL; AJ240806; CAB53480.1; -  
 DR EMBL; AJ240807; CAB53481.1; -  
 DR EMBL; AJ240808; CAB53482.1; -  
 DR EMBL; AJ240809; CAB53483.1; -  
 DR EMBL; AJ240810; CAB53484.1; -  
 DR EMBL; AJ240811; CAB53485.1; -  
 DR EMBL; AJ240812; CAB53486.1; -  
 DR EMBL; AJ240813; CAB53487.1; -  
 DR EMBL; AJ240814; CAB53488.1; -  
 DR EMBL; AJ240815; CAB53489.1; -  
 DR EMBL; AJ240816; CAB53490.1; -  
 DR EMBL; AJ240817; CAB53491.1; -  
 DR EMBL; AJ240818; CAB53492.1; -  
 DR EMBL; AJ240819; CAB53493.1; -  
 DR EMBL; AJ240820; CAB53494.1; -  
 DR EMBL; AJ240821; CAB53495.1; -  
 DR EMBL; AJ240822; CAB53496.1; -  
 DR EMBL; AJ240823; CAB53497.1; -  
 DR EMBL; AJ240824; CAB53498.1; -  
 DR EMBL; AJ240825; CAB53499.1; -  
 DR EMBL; AJ240826; CAB53500.1; -  
 DR EMBL; AJ240827; CAB53501.1; -  
 DR EMBL; AJ240828; CAB53502.1; -  
 DR EMBL; AJ240829; CAB53503.1; -  
 DR EMBL; AJ240830; CAB53504.1; -  
 DR EMBL; AJ240831; CAB53505.1; -  
 DR EMBL; AJ240832; CAB53506.1; -  
 DR EMBL; AJ240833; CAB53507.1; -  
 DR EMBL; AJ240834; CAB53508.1; -  
 DR EMBL; AJ240835; CAB53509.1; -  
 DR EMBL; AJ240836; CAB53510.1; -  
 DR EMBL; AJ240837; CAB53511.1; -  
 DR EMBL; AJ240838; CAB53512.1; -  
 DR EMBL; AJ240839; CAB53513.1; -  
 DR EMBL; AJ240840; CAB53514.1; -  
 DR EMBL; AJ240841; CAB53515.1; -  
 DR EMBL; AJ240842; CAB53516.1; -  
 DR EMBL; AJ240843; CAB53517.1; -  
 DR EMBL; AJ240844; CAB53518.1; -  
 DR EMBL; AJ240845; CAB53519.1; -  
 DR EMBL; AJ240846; CAB53520.1; -  
 DR EMBL; AJ240847; CAB53521.1; -  
 DR EMBL; AJ240848; CAB53522.1; -  
 DR EMBL; AJ240849; CAB53523.1; -  
 DR EMBL; AJ240850; CAB53524.1; -  
 DR EMBL; AJ240851; CAB53525.1; -  
 DR EMBL; AJ240852; CAB53526.1; -  
 DR EMBL; AJ240853; CAB53527.1; -  
 DR EMBL; AJ240854; CAB53528.1; -  
 DR EMBL; AJ240855; CAB53529.1; -  
 DR EMBL; AJ240856; CAB53530.1; -  
 DR EMBL; AJ240857; CAB53531.1; -  
 DR EMBL; AJ240858; CAB53532.1; -  
 DR EMBL; AJ240859; CAB53533.1; -  
 DR EMBL; AJ240860; CAB53534.1; -  
 DR EMBL; AJ240861; CAB53535.1; -  
 DR EMBL; AJ240862; CAB53536.1; -  
 DR EMBL; AJ240863; CAB53537.1; -  
 DR EMBL; AJ240864; CAB53538.1; -  
 DR EMBL; AJ240865; CAB5353

[illegible]



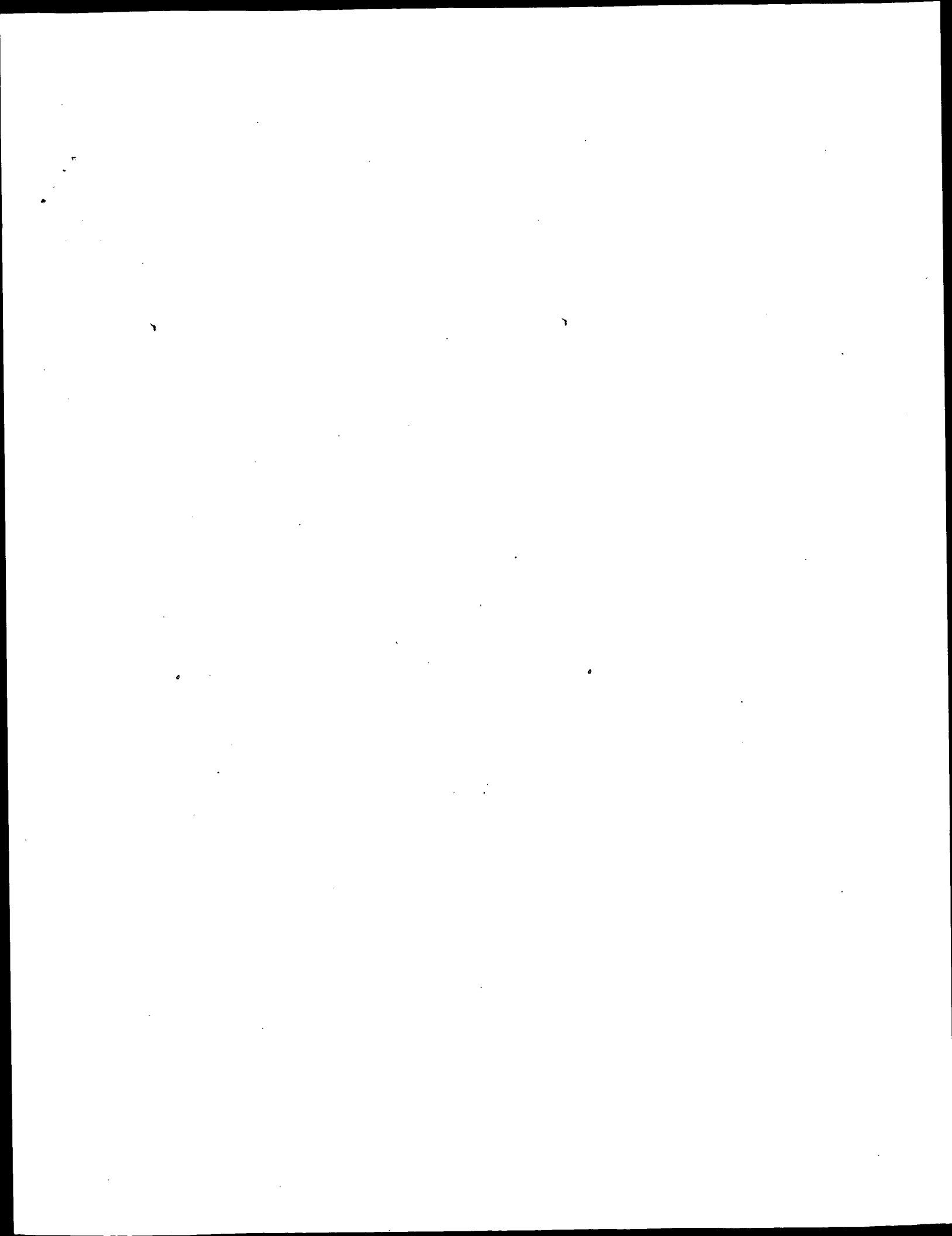
15-JUN-1998 (Rel. 36, Last annotation update)  
 NMDH-plastoquinone oxidoreductase chain 5, chloroplast (EC 1.6.5.3)  
 (fragment).  
 DE NMDH.  
 GN NDHF.  
 OS Digitalis grandiflora (Yellow foxglove).  
 OG Chloroplast.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; euasterids I; Lamiales; Veronicaceae; Digitalis.  
 CC NCBI\_TaxID=38791;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Olmstead R.G., Reeves P.A.;  
 RT "Evidence for the polyphyly of the Scrophulariaceae based on  
 chloroplast rbcL and ndhF sequences." f  
 RL Ann. Mo. Bot. Gard. 82:176-193(1995).  
 CC -1- CATALYTIC ACTIVITY: NMDH + plastoquinone = NMD(+) + plastoquinol.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L36399; AAA84203.1; -  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR InterPro; IPR002128; Oxidored\_q1\_C.  
 DR InterPro; IPR001516; Oxidored\_q1\_N.  
 DR Pfam; PF00361; Oxidored\_q1\_1.  
 DR Pfam; PF00662; Oxidored\_q1\_N; 1.  
 DR Pfam; PF01010; Oxidored\_q1\_C; 1.  
 DR Oxidoreductase; NAD; Plastoquinone; Chloroplast.  
 FT NON TER 1 1  
 FT NON TER 699 699  
 SQ SEQUENCE 699 AA; 79309 MW; 5685FBA561C63D01 CRC64;  
 Query Match 11.6%; Score 65.5; DB 1; Length 699;  
 Best Local Similarity 30.3%; Pred. No. 7.8;  
 Matches 23; Conservative 9; Mismatches 25; Indels 19; Gaps 3;  
 QY 33 IPAS-KNFILOKNTIGTGR-----MKNNIHWLOFAKLTGPTLMGRRLKMP 78  
 DB 618 IPASFENFDLINSFVYTGKRSRWDKILTLNWSHNPAYIDVYPTSTGSIKLSOLT 677  
 QY 79 YIAGYKTFDGRVDGI 94  
 DB 678 HF-----FDYQVIDGI 688  
 RESULT 14  
 AMR DROR STANDARD; PRT; 493 AA.  
 AC 077015;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Alpha-amyase-related protein precursor (EC 3.2.1.1).  
 GN AMYREL.  
 OS Drosophila arena (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7233;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA De Laje J.-L.;  
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO

KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U96158; AAC39108.2; -  
 DR HSSP; P56634; IJAE.  
 DR Flybase; FBgn0021266; Dore\Amyrel.  
 DR InterPro; IPR000461; Alpha\_amyase.  
 DR Pfam; PF00128; alpha-amyase; 1.  
 DR Pfam; PF02806; alpha-amyase\_C; 1.  
 DR PRINTS; PR00110; ALPHAMYLASE.  
 KM Hydrolase; Glycosidase; Carbohydrate metabolism; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 493  
 FT ACT\_SITE 207 207  
 FT ACT\_SITE 211 211  
 FT ACT\_SITE 309 309  
 FT ACT\_SITE 47 103  
 FT DISULFID 156 170  
 FT DISULFID 417 440  
 FT DISULFID 447 459  
 FT DISULFID 493 553  
 SQ SEQUENCE 493 AA; 55384 MW; 87BE4183B3F15E0 CRC64;  
 Query Match 11.6%; Score 64.5; DB 1; Length 493;  
 Best Local Similarity 28.7%; Pred. No. 6.9;  
 Matches 31; Conservative 10; Mismatches 42; Indels 25; Gaps 6;  
 QY 10 FGNRRVFIRKSVNVIHGRI-----DIFASKNFILOKNTIGTGRMKNNIHWLO-F 61  
 DB 233 FPHNSRPF--FQEVYIDHGHETVSRSDEYKDLGAVTFEPRSEE-IGNAFRGNNAKMLQSW 289  
 QY 62 AKLTGFTLMGRRLKMPYIAGYKTF-----DGRVDGIITAAVONPASK 105  
 DB 290 GTGCGFLPBGQAL-----TFVDNHNORDAGAVLSKSPKPYK 327  
 RESULT 15  
 RS9 NEIMA STANDARD; PRT; 130 AA.  
 ID RS9 NEIMA  
 AC 090Q29;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 30S ribosomal protein S9.  
 GN RPSI OR NMA0379 OR NMB2056.  
 OS Neisseria meningitidis (serogroup A), and  
 OS Neisseria meningitidis (serogroup B).  
 CC Bacteria; Proteobacteria; Beta subdivision; Neisseriaceae; Neisseria.  
 CC NCBI\_TaxID=65699, 491;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Z2491 / Serogroup A / Serotype 4A;  
 RX MEDLINE=20222556; PubMed=10761919;  
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
 RA Davies R.M., Davis P., Devlin K., Feltham T., Hamlin N., Holtroyd S.,  
 RA Jørgensen K., Leather S., Moule S., Mungall K., Quail M.A.,  
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
 RA Whitehead S., Spratt B.G., Barrall B.G.;  
 RT "Complete DNA sequence of a serogroup A strain of Neisseria  
 meningitidis Z2491.";  
 RT Nature 404:502-506(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=MC58 / Serogroup B;  
 RX MEDLINE=20157555; PubMed=10710307;  
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,



Search completed: April 20, 2003, 13:07:28  
Job time : 23.4474 secs

Search completed: April 20, 2003, 13:07:28  
Job time : 23.4474 secs



GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:56:55 ; Search time 88.6974 Seconds  
(without alignments)  
243.919 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 538  
Sequence: 1 MKVTVAFNQFGPNRRVPIKR.....FDGRVDGIIAAYGNPASWK 105

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: SP\_archaea:\*  
2: SP\_bacteria:\*  
3: SP\_fungi:\*  
4: SP\_human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mmc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_rodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_virus:\*  
16: SP\_bacterioph:\*  
17: SP\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93.5	16.8	374	10 Q8RUR1	Q8RUR1 cryptomeria
2	91.5	16.4	507	10 Q9FY19	Q9FY19 juniperus a
3	86.5	15.5	367	10 Q9XIL6	Q9XIL6 cupressus a
4	85.5	15.3	367	10 Q9LUT2	Q9LUT2 juniperus v
5	85.5	15.3	367	10 Q9LUT1	Q9LUT1 juniperus v
6	85.5	15.3	367	10 Q9MAS6	Q9MAS6 cupressus s
7	85.5	15.3	367	10 Q9MAS4	Q9MAS4 cupressus s
8	85.5	15.3	367	10 Q9MAS3	Q9MAS3 cupressus s
9	85.5	15.3	367	10 Q9MAS2	Q9MAS2 cupressus s
10	85.5	15.3	367	10 Q9XSL1	Q9XSL1 juniperus o
11	84.5	15.1	398	16 Q9WZL9	Q9WZL9 thermotoga
12	84.5	14.6	367	10 Q9MAS5	Q9MAS5 cupressus s
13	81	14.5	219	10 Q9SIO9	Q9SIO9 arabidopsis
14	77	13.8	219	10 Q9J377	Q9J377 arabidopsis
15	74	13.3	221	10 Q9LWK1	Q9LWK1 oryza sativ
16	73.5	13.2	644	16 Q8YEX3	Q8YEX3 brucea me

17	72.5	13.0	440	10 Q949D3	Q949D3 oryza sativ
18	71.5	12.8	645	16 Q96SQ4	Q96SQ4 rhizobium 1
19	71.5	12.8	700	8 Q9THR5	Q9THR5 calceolaria
20	68.5	12.3	390	16 Q9THB6	Q9THB6 clostridium
21	68	12.2	167	2 Q9KHV1	Q9KHV1 bordetella
22	68	12.2	199	2 Q8VP01	Q8VP01 bordetella
23	68	12.2	221	10 Q9LFR1	Q9LFR1 arabidopsis
24	68	12.2	576	10 Q9LV91	Q9LV91 arabidopsis
25	68	12.2	690	11 Q9ES38	Q9ES38 ratius norv
26	67	12.0	572	8 Q9TII7	Q9TII7 mauiensis p
27	67	12.0	662	11 Q88694	Q88694 mus musculu
28	67	12.0	689	11 Q91VDS	Q91VDS mus musculu
29	67	12.0	695	8 Q9TK54	Q9TK54 talium ang
30	66.5	11.9	493	5 Q9BH28	Q9BH28 dirosophila
31	66.5	11.9	493	5 Q9BN06	Q9BN06 dirosophila
32	66.5	11.9	686	8 P92329	P92329 plectranthu
33	66.5	11.9	746	8 Q9TLB7	Q9TLB7 digitalis p
34	66.5	11.9	747	8 Q9TLA7	Q9TLA7 palisana p
35	66.5	11.9	748	8 Q94VZ8	Q94VZ8 pyralia 1
36	66.5	11.9	777	17 Q9J763	Q9J763 archaeoglob
37	66.5	11.9	1429	5 Q960A1	Q960A1 caenorhabdi
38	66	11.8	188	17 Q9Y9P2	Q9Y9P2 aeropyrum p
39	66	11.8	745	8 Q9TLC1	Q9TLC1 clarendoru
40	65.5	11.7	602	16 Q8U966	Q8U966 agrobacteri
41	65.5	11.7	699	8 Q9TLF3	Q9TLF3 hemipnagma
42	65.5	11.7	700	8 Q9THN4	Q9THN4 joveilana s
43	65.5	11.7	750	16 Q9CNC8	Q9CNC8 pasteurilla
44	65	11.6	280	2 Q52104	Q52104 ethlichia c
45	65	11.6	335	2 Q9AHF2	Q9AHF2 agrobacteri

## ALIGNMENTS

RESULT 1	Q8RUR1	PRELIMINARY;	PRT;	374 AA.
ID	Q8RUR1			
AC	Q8RUR1;			
DT	01-JUN-2002 (TEMBLrel. 21, Created)			
DT	01-JUN-2002 (TEMBLrel. 21, Last sequence update)			
DT	01-JUN-2002 (TEMBLrel. 21, Last annotation update)			
DE	Cry j 1 precursor.			
GN	Cry j 1.1 OR Cry j 1.2.			
OS	Cryptomeria japonica (Japanese cedar).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.			
OX	NCBI_TaxID=3369;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISUE=POLEEN;			
RA	Futamura N., Shinohara K.,			
RT	"Isolation and characterization of cDNAs encoding major allergen Cry j			
RT	1 from Cryptomeria japonica pollen."			
RL	Submitted (Mar-2002) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AB081309; BAB86286.1; -			
DR	EMBL; AB081310; BAB86287.1; -			
KW	Signal.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	374	Cry j 1.
FT	SEQUENCE	374 AA;	40720 MW;	90D0085D24BF2BD4 CRC64;
QY	Query Match	16.8%;	Score 93.5;	DB 10; Length 374;
QY	Best Local Similarity	51.4%;	Pred. No. 0.017;	7; Indels 5; Gaps 1;
QY	Matches 19;	Conservative 6;	Mismatches 7;	
QY	53 KNNRWLPQAKLTGFTLMGRRLKMPYIAGYTFDGR 89			
QY	85 RDRPLWIRFS-----GNMWRKMKMPYIAGYTFDGR 116			
RESULT 2	Q9FY19	PRELIMINARY;	PRT;	507 AA.
ID	Q9FY19			

AC Q9FY19; 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Pollen major allergen 2 protein precursor.  
 GN JNA2.  
 OS Juniperus ashei (Ozark white cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 NCBI\_TaxID=1101;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MALE POLLEN;  
 RX MEDLINE=20403896; PubMed=10944464;  
 RA Yokoyama M., Miyahara M., Shimizu K., Kino K., Tsunoo H.;  
 RT "Purification, identification and cDNA cloning of Jun a 2, the second  
 major allergen of mountain cedar pollen."  
 RL Biochem Biophys Res Commun 275:195-202(2000).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES  
 (POLYGALACTURONASES).  
 DR EMBL; AJ404653; CAC05582.1; -.  
 DR HSSP; P26509; 1BHE.  
 DR InterPro; IPR000743; GH28.  
 DR InterPro; IPR000408; Reg chr condens.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR PROSITE; PS00502; POLYGALACTURONASE; UNKNOWN\_1.  
 DR PROSITE; PS00626; RCO1\_2; UNKNOWN\_1.  
 KW Cell wall; Glycosidase; Hydrolase; Signal.  
 FT SIGNAL 1 54 POTENTIAL.  
 SQ SEQUENCE 507 AA; 55730 MW; 2B2E0A5E5958FE5A CRC64;  
 Query Match 16.4%; Score 91.5; DB 10; Length 507;  
 Best Local Similarity 31.5%; Pred. No. 0.041;  
 Matches 29; Conservative 14; Mismatches 26; Indels 23; Gaps 5;  
 QY 12 ENRRVYIKSVNVIHGRIDIFASGNFLQNTIGT---GRMNKNRIWOPAKLTGFT 68  
 DB 97 PANKFTV---VNNLVYGGPCQPHFSFK---VDGTTAAVDPKAKWNSKIMHFAKLTDFN 150  
 QY 69 LMGRLKMPYIAGY-----KTFDGRV 91  
 DB 151 LMGTV-----IDGQNRWMSDCKTINRTV 177  
 RESULT 3  
 Q93XL6 PRELIMINARY; PRT; 367 AA.  
 AC Q93XL6; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative allergen Cup a 1 precursor.  
 GN CUP A 1.  
 OS Cupressus arizonica.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 NCBI\_TaxID=49011;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=POLLEN;  
 RA Buteront C., Di Felice G., Pini C.;  
 RT "Cloning of Cupressus arizonica major allergen."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBS databases.  
 DR EMBL; A0278498; CAC37790.2; -.  
 KW Signal.  
 FT CHAIN 1 21 POTENTIAL.  
 FT SIGNAL 22 367 PUTATIVE ALLERGEN CUP A 1.  
 SQ SEQUENCE 367 AA; 39809 MW; AFE97260423A9F28 CRC64;  
 Query Match 15.5%; Score 86.5; DB 10; Length 367;  
 Best Local Similarity 40.0%; Pred. No. 0.11;  
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGRVD 92  
 DB 85 REKALWIIFSQ-----NMNKKLMPLYVAGHKTIDGRAD 119  
 RESULT 4  
 Q9LTL2 PRELIMINARY; PRT; 367 AA.  
 AC Q9LTL2; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Pollen major allergen 1.2.  
 OS Juniperus virginiana (Eastern red cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 NCBI\_TaxID=39584;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21315424; PubMed=11422137;  
 RA Midoro-Horiuti T.M., Goldblum R.M., Brooks E.G.;  
 RT "Identification of mutations in the genes for the pollen allergens of  
 eastern red cedar (Juniperus virginiana)."  
 RL Clin. Exp. Allergy 31:771-778(2001).  
 DR EMBL; AF151427; AAF80164.1; -.  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PRO0807; AMBALLERGEN.  
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;  
 Query Match 15.3%; Score 85.5; DB 10; Length 367;  
 Best Local Similarity 40.0%; Pred. No. 0.14;  
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;  
 QY 53 KNNRIWLOFAKLTGFTLMGRRLKMPYIAGYKTFDGRVD 92  
 DB 85 REKALWIIFSQ-----NMNKKLMPLYVAGHKTIDGRAD 119  
 RESULT 5  
 Q9LTL1 PRELIMINARY; PRT; 367 AA.  
 AC Q9LTL1; 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Pollen major allergen 1-1.  
 OS Juniperus virginiana (Eastern red cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 NCBI\_TaxID=39584;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21315424; PubMed=11422137;  
 RA Midoro-Horiuti T.M., Goldblum R.M., Brooks E.G.;  
 RT "Identification of mutations in the genes for the pollen allergens of  
 eastern red cedar (Juniperus virginiana)."  
 RL Clin. Exp. Allergy 31:771-778(2001).  
 DR EMBL; AF151429; AAF80166.1; -.  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PRO0807; AMBALLERGEN.  
 SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;  
 Query Match 15.3%; Score 85.5; DB 10; Length 367;  
 Best Local Similarity 40.0%; Pred. No. 0.14;  
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

```

RESULT 6
QY 53 KNNRIMLOPAKLTGFTLMGRRLKMPYIAGYTFDGRVD 92
ID Q9M4S6 PRELIMINARY; PRT; 367 AA.
AC Q9M4S6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257491; AAF72625.1;
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; B4E9C60108C2C5A3 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIMLOPAKLTGFTLMGRRLKMPYIAGYTFDGRVD 92
ID Q9M4S4 PRELIMINARY; PRT; 367 AA.
AC Q9M4S4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257493; AAF72627.1;
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIMLOPAKLTGFTLMGRRLKMPYIAGYTFDGRVD 92
ID Q9M4S3 PRELIMINARY; PRT; 367 AA.
AC Q9M4S3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)

```

```

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257494; AAF72628.1;
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39832 MW; B5DFBF5A61C07A53 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIMLOPAKLTGFTLMGRRLKMPYIAGYTFDGRVD 92
ID Q9M4S2 PRELIMINARY; PRT; 367 AA.
AC Q9M4S2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP51.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF257495; AAF72629.1;
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39819 MW; AE7E055A61C07A53 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;
Best Local Similarity 40.0%; Pred. No. 0.14;
Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

QY 53 KNNRIMLOPAKLTGFTLMGRRLKMPYIAGYTFDGRVD 92
ID Q93X51 PRELIMINARY; PRT; 367 AA.
AC Q93X51;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Putative allergen Jun o 1.
GN JUN O 1.
OS Juniperus oxycedrus (Prickly juniper).

```

OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 OX NCBI\_TaxId=69008;  
 RN (1)  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=POLLEN;  
 RA Iacovacci P., Di Felice G., Plant C.;  
 RT "Cloning of Juniperus oxycedrus major allergen."  
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ293767; CAC48400.1; -  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBFD1B9D7 CRC64;

Query Match 15.3%; Score 85.5; DB 10; Length 367;  
 Best Local Similarity 40.0%; Pred. No. 0.14;  
 Matches 16; Conservative 9; Mismatches 10; Indels 5; Gaps 1;

DB 53 KNNRWLOPAKLTGFTLMGRRLKMPYIAGYKTFPGRRVD 92  
 85 REKALMIIFISQ-----NMNIKLMPLVAGHKTIDRGAD 119

RESULT 11

O9WZL9 PRELIMINARY; PRT; 398 AA.  
 AC O9WZL9;  
 DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Lipopolysaccharide biosynthesis protein, putative.  
 GN TM0760.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.  
 OX NCBI\_TaxId=2336;  
 RN (1)  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=MSB8 / DSM 3109;  
 RX MEDLINE=99287316; PubMed=10360571;  
 RA Hafe D.H., Clayton R.A., Gill S.R., Grimm M.L., Dodson R.J.,  
 RA McDonald L., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eissen J.A., White O.,  
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
 genome sequence of Thermotoga maritima."  
 RL Nature 399:323-329 (1999).  
 DR EMBL; AE001746; AAD35842.1; -  
 DR TIGR; TM0760; -  
 DR InterPro; IPR001296; Glycos\_transf\_1.  
 DR Pfam; PF00534; Glycos\_transf\_1; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 398 AA; 46448 MW; 28C5F84D55A25912 CRC64;

Query Match 15.1%; Score 84.5; DB 16; Length 398;  
 Best Local Similarity 28.4%; Pred. No. 0.2;  
 Matches 25; Conservative 11; Mismatches 35; Indels 17; Gaps 2;

OY 28 GRRIDIFASKNFHLQKNTGTGRKNNRIW-----LQFAKLTGFTLMGRRLK 76  
 DB 33 GHRVDIYIGDFSHL-----TGKRWSETFGWSFKDGVDFIVERREYTGSLRLSSI 86  
 OY 77 PVIAGYKTFPGRRVDGIIAQNPAW 104  
 DB 87 DYRNKRKKTIOKRYDVIIASSPHPSW 114

RESULT 12

O9M4S5 PRELIMINARY; PRT; 367 AA.  
 AC O9M4S5;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Cup s 1 pollen allergen.  
 GN CUPSL.  
 OS Cupressus sempervirens.  
 OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 OX NCBI\_TaxId=13469;  
 RN (1)  
 RC SEQUENCE FROM N.A.  
 RP Monsalve R.I., Villalba M., Rodriguez R.;  
 RT "Cloning and expression of Cup s 1, the major allergen of the pollen  
 of Cupressus sempervirens."  
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF257492; AAF72626.1; -  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PR00807; AMBALLERGEN.  
 SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match 14.6%; Score 81.5; DB 10; Length 367;  
 Best Local Similarity 37.5%; Pred. No. 0.4;  
 Matches 15; Conservative 10; Mismatches 10; Indels 5; Gaps 1;

OY 53 KNNRWLOPAKLTGFTLMGRRLKMPYIAGYKTFPGRRVD 92  
 85 REKALMIIFISQ-----NMNIKLMPLVAGHKTIDRGAD 119

RESULT 13

O9S1O9 PRELIMINARY; PRT; 219 AA.  
 AC O9S1O9;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Putative synaptobrevin.  
 GN AT025340.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxId=3702;  
 RN (1)  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buehl C.R., Ketchum K.A., Lee J.U., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Coppenhaver G.P., Preuss D., Niernan W.C., White O., Eissen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana."  
 RL Nature 402:761-768 (1999).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC007070; AAD3657.1; -  
 DR InterPro; IPR001388; Synaptobrevin.  
 DR Pfam; PF00957; synaptobrevin; 1.  
 DR PRINTS; PR00219; SYNAPTOBREVN.  
 DR ProDom; PD001229; Synaptobrevin; 1.  
 SQ SEQUENCE 219 AA; 24965 MW; 7A63C85A140913B9 CRC64;

Query Match 14.5%; Score 81; DB 10; Length 219;  
 Best Local Similarity 29.6%; Pred. No. 0.24;  
 Matches 16; Conservative 14; Mismatches 22; Indels 2; Gaps 1;

OY 14 RRVFIRKRVSNVTHGRRIDIFASKNFHLOKNTI---GTGRBKNRRIWLOFAKLT 65  
 DB 138 RGVMIENIDKVLDRGERLELVDTKANNQGNIFRFRKQTRFRFNNVTWVRNCKLT 191

## RESULT 14

OY 049377 PRELIMINARY; PRT; 219 AA.  
 AC 049377;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DB Synapcobrevin-like protein (AT4G32150/Fl0N7\_40).  
 GN Fl0N7\_40 OR AT VAMP7C OR AT4G32150.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 CX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Beyan M., Koeltter P., Hempel S., Entian K.-D., Hobeisel J.,  
 RA Mewes H.W., Mayer K.F.X., Scheller C.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Nikoloff D.M., Somerville C.R.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,  
 RA Bahr J., Bower L., Carninci P., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamuya A.,  
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,  
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis cDNA clones."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021636; CAAL6574.1; -  
 DR EMBL; AF025332; AAD01748.1; -  
 DR EMBL; AL161580; CAB79933.1; -  
 DR EMBL; AF439840; AAL27509.1; -  
 DR InterPro; IPR001388; Synapcobrevin.  
 DR Pfam; PF00957; synapcobrevin; 1.  
 DR PRINTS; PR00219; SYNAPTOBREVN.  
 DR PRODOM; PD001229; Synapcobrevin; 1.  
 SQ SEQUENCE 219 AA; 25039 MW; 54C3BGF5C7A3D39E CRC64;

Query Match 13.8%; Score 77; DB 10; Length 219;  
 Best Local Similarity 31.0%; Pred. No. 0.7;  
 Matches 22; Conservative 16; Mismatches 27; Indels 6; Gaps 3;

OY 14 RRVFIRKRVSNVTHGRRIDIFASKNFHLOKNTI---GTGRBKNRRIWLOFAKLT 70  
 DB 138 RGVMIENIDKVLDRGERLELVDTKANNQGNIFRFRKQTRFRFNNVTWVRNCKLT--VLL 194  
 OY 71 GRRLKMPVYA 81  
 DB 195 ILLLVITYIA 205

RESULT 15  
 O9LWK1

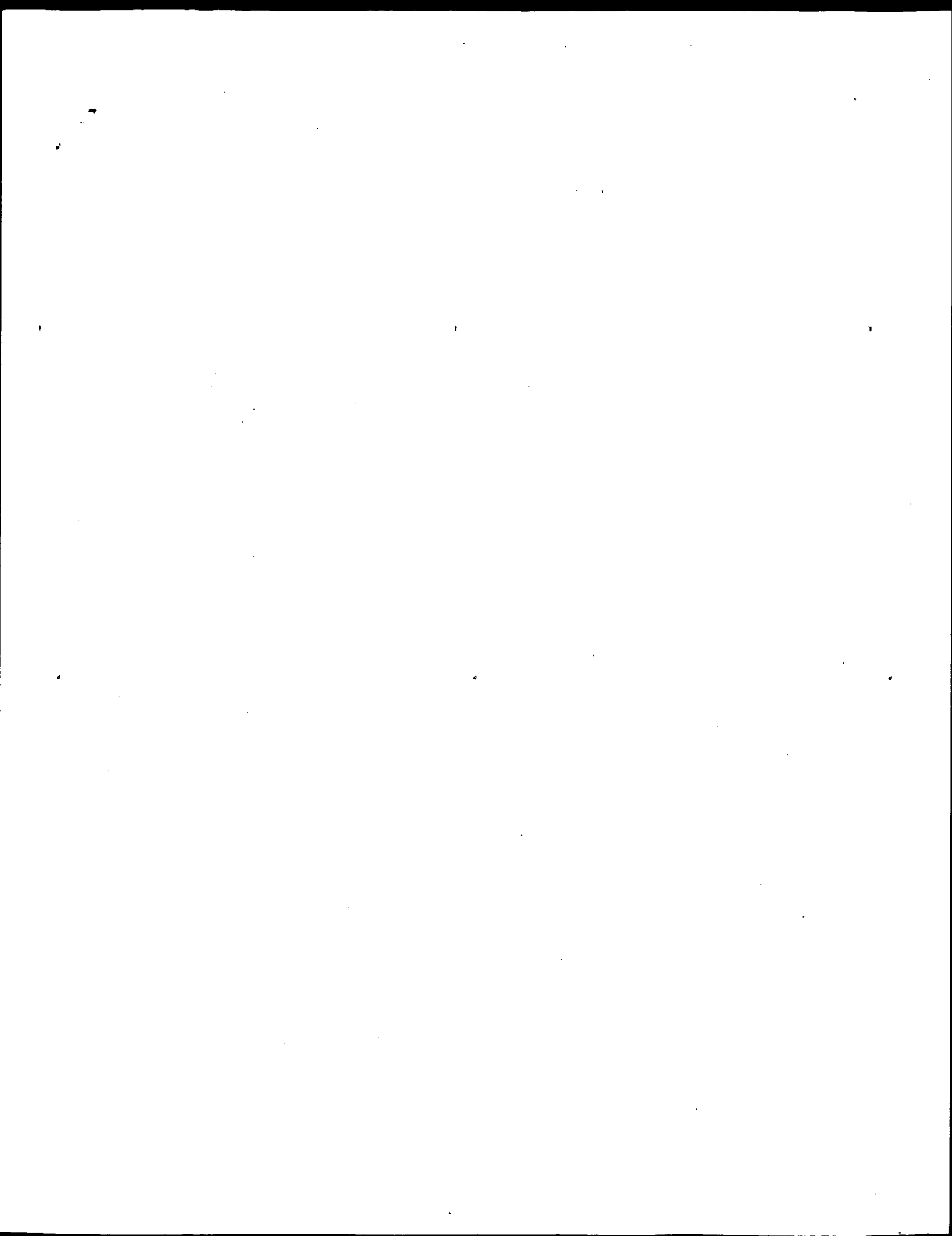
ID 09LWK1 PRELIMINARY; PRT; 221 AA.

AC 09LWK1;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE ESTs AU082579(S2069).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae;  
 OC Eriactoidae; Oryzae; Oryza.  
 CX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC  
 RT clone: P0015E04."  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP002069; BAA95814.1; -  
 DR InterPro; IPR001388; Synapcobrevin.  
 DR Pfam; PF00957; synapcobrevin; 1.  
 DR PRINTS; PR00219; SYNAPTOBREVN.  
 DR PRODOM; PD001229; Synapcobrevin; 1.  
 SQ SEQUENCE 221 AA; 24970 MW; 1D573364B8E38571 CRC64;

Query Match 13.3%; Score 74; DB 10; Length 221;  
 Best Local Similarity 29.1%; Pred. No. 1.6;  
 Matches 23; Conservative 9; Mismatches 23; Indels 24; Gaps 3;

OY 14 RRVFIRKRVSNVTHGRRIDIFASKNFHLOKNTI---GTGRBKNRRIWLOFAKLT 65  
 DB 139 RTVMINIDKVLDRGERLELVDTKANNQGNIFRFRKQTRFRFNNVTWVRNCKLTALIF 197  
 OY 66 -----GFTL 69  
 DB 198 LTVIIVVIVPMCHGFTL 216

Search completed: April 20, 2003, 13:12:57  
 Job time : 91.6974 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:05 ; Search time 127.382 Seconds  
(without alignments)  
109.838 Million cell updates/sec

Title: US-09-142-524D-2

Perfect score: 558  
Sequence: 1 MKVTVAFNQFGNRRVFIK.....FDGRVDGIIAAYNPASMK 105

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A\_Geneseq\_101002:\*  
2: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
24: /SIDB2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558	100.0	105	AAW27370	Multi-epitope pept
2	523.5	93.8	134	AAW27371	Multi-epitope pept
3	327.5	58.7	80	AAW27369	Multi-epitope pept
4	213	38.2	214	AAW27369	Cedar pollen aller
5	192.5	34.5	210	AAW27369	Cedar pollen aller
6	162	29.0	96	AAW27369	Cedar pollen aller
7	162	29.0	96	AAW27369	Cedar pollen aller
8	162	29.0	96	AAW27369	Cedar pollen aller
9	149.5	26.8	97	AAW27369	Cedar pollen aller
10	141.5	25.4	93	AAW27369	Cedar pollen aller

11	131.5	23.6	47	19	AAW80353	Sugi allergen prot
12	120.5	21.6	74	21	AAW23905	Artificial sequenc
13	120.5	21.6	94	21	AAW23906	Artificial sequenc
14	120.5	21.6	95	21	AAW23907	Artificial sequenc
15	120.5	21.6	95	21	AAW23908	Artificial sequenc
16	120.5	21.6	95	21	AAW23909	Artificial sequenc
17	120.5	21.6	95	21	AAW23910	Artificial sequenc
18	120.5	21.6	95	21	AAW23911	Artificial sequenc
19	120.5	21.6	95	21	AAW23912	Artificial sequenc
20	115.5	20.7	514	16	AAW23913	Artificial sequenc
21	115.5	20.7	514	16	AAW23914	Artificial sequenc
22	115.5	20.7	514	16	AAW23915	Artificial sequenc
23	115	20.6	514	20	AAW23916	Artificial sequenc
24	115	20.6	514	16	AAW23917	Artificial sequenc
25	115	20.6	514	16	AAW23918	Artificial sequenc
26	115	20.6	514	16	AAW23919	Artificial sequenc
27	115	20.6	514	20	AAW23920	Artificial sequenc
28	106.5	19.1	71	19	AAW23921	Artificial sequenc
29	106.5	19.1	81	21	AAW23922	Artificial sequenc
30	106.5	19.1	81	21	AAW23923	Artificial sequenc
31	106.5	19.1	81	21	AAW23924	Artificial sequenc
32	106.5	19.1	81	21	AAW23925	Artificial sequenc
33	106.5	19.1	81	21	AAW23926	Artificial sequenc
34	106.5	19.1	81	21	AAW23927	Artificial sequenc
35	106.5	19.1	81	21	AAW23928	Artificial sequenc
36	106.5	19.1	81	21	AAW23929	Artificial sequenc
37	106.5	19.1	81	21	AAW23930	Artificial sequenc
38	106.5	19.1	81	21	AAW23931	Artificial sequenc
39	103	18.5	42	22	AAW23932	Artificial sequenc
40	103	18.5	42	22	AAW23933	Artificial sequenc
41	101.5	18.2	47	19	AAW23934	Artificial sequenc
42	101	18.1	47	19	AAW23935	Artificial sequenc
43	99	17.7	19	22	AAW23936	Artificial sequenc
44	99	17.7	19	22	AAW23937	Artificial sequenc
45	99	17.7	19	22	AAW23938	Artificial sequenc

## ALIGNMENTS

RESULT 1	AAW27370	standard; peptide; 105 AA.
ID	AAW27370	
XX	AAW27370:	
AC	AAW27370:	
XX	AAW27370:	
DT*	24-MAR-1998 (first entry)	
XX	Multi-epitope peptide used as immunotherapeutic agent #2.	
DE	Multi-epitope peptide; immunotherapeutic agent; allergic disease;	
XX	T-cell epitope region; allergen; lymphocyte; immunoglobulin E.	
KW	Synthetic.	
OS	MO9732600-AA.	
XX	MO9732600-AA.	
PN	12-SEP-1997.	
XX	12-SEP-1997.	
PD	10-MAR-1997; 97WO-UP00740.	
XX	10-MAR-1997; 97WO-UP00740.	
PF	10-MAR-1996; 96UP-0080702.	
XX	10-MAR-1996; 96UP-0080702.	
PR	(MEIP) MEIJI MILK PROD CO LTD.	
XX	(MEIP) MEIJI MILK PROD CO LTD.	
PA	Dairiki K, Iwama A, Kino K, Kume A, Sone T;	
XX	Dairiki K, Iwama A, Kino K, Kume A, Sone T;	
XX	WPI; 1997-470495/43.	
DR	Peptide immuno-therapeutic agent to treat allergic diseases -	
XX	contains multi-epitope peptide containing T cell epitope regions	
PT	from different allergens	
XX		

PS Claim 6; Page 31; 58bp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 105 AA;

Query Match 100.0%; Score 558; DB 18; Length 105;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-65;  
 Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKVTAFNPGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRKNNRIWLQ 60  
 |||  
 Db 1 MKVTAFNPGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRKNNRIWLQ 60  
 |||

Qy 61 FAKLTGFTLMGRRLKMPYIAGYKTFDGRVDTIAYQNPASMK 105  
 |||  
 Db 61 FAKLTGFTLMGRRLKMPYIAGYKTFDGRVDTIAYQNPASMK 105  
 |||

RESULT 2  
 AAW27371  
 ID AAW27371 standard; peptide; 134 AA.

AC AAW27371;  
 XX  
 DT 24-MAR-1998 (first entry)

XX Multi-epitope peptide used as immunotherapeutic agent #3.  
 DE  
 XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 OS Synthetic.  
 XX  
 PN WO9732600-A1.  
 XX  
 PD 12-SEP-1997.  
 XX  
 PF 10-MAR-1997; 97WO-JP00740.  
 XX  
 PR 10-MAR-1996; 96JP-0080702.  
 XX  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 XX  
 PI Daijiki K, Iwama A, Kuno K, Kume A, Sone T;  
 XX  
 DR WPI; 1997-470495/43.  
 XX  
 PT Peptide immuno:therapeutic agent to treat allergic diseases -  
 PT contains multi-epitope peptide containing T cell epitope regions  
 PT from different allergens  
 XX  
 PS Claim 6; Page 32; 58bp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

XX Sequence 134 AA;

SQ  
 Query Match 93.8%; Score 523.5; DB 18; Length 134;  
 Best Local Similarity 78.4%; Pred. No. 1.4e-60;  
 Matches 105; Conservative 0; Mismatches 0; Indels 29; Gaps 2;

Qy 1 MKVTAFNPGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRKNNRIWLQ 60  
 |||  
 Db 1 MKVTAFNPGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRKNNRIWLQ 60  
 |||

Qy 61 FAKLTGFTLMGR-----LKMPTIAGYKTFD-----RRV 91  
 |||  
 Db 61 FAKLTGFTLMGRPLMIIFSGNMNIKMPYIAGYKTFDGRRAVSYVHNGAKFTIRV 120  
 |||

Qy 92 DGIHAYQNPASMK 105  
 |||  
 Db 121 DGIHAYQNPASMK 134  
 |||

RESULT 3  
 AAW27369  
 ID AAW27369 standard; peptide; 80 AA.

AC AAW27369;  
 XX  
 DT 24-MAR-1998 (first entry)

XX Multi-epitope peptide used as immunotherapeutic agent #1.  
 DE  
 XX Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 OS Synthetic.  
 XX  
 PN WO9732600-A1.  
 XX  
 PD 12-SEP-1997.  
 XX  
 PF 10-MAR-1997; 97WO-JP00740.  
 XX  
 PR 10-MAR-1996; 96JP-0080702.  
 XX  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 XX  
 PI Daijiki K, Iwama A, Kuno K, Kume A, Sone T;  
 XX  
 DR WPI; 1997-470495/43.  
 XX  
 PT Peptide immuno:therapeutic agent to treat allergic diseases -  
 PT contains multi-epitope peptide containing T cell epitope regions  
 PT from different allergens  
 XX  
 PS Claim 6; Page 31; 58bp; Japanese.

XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 80 AA;

Query Match 58.7%; Score 327.5; DB 18; Length 80;  
 Best Local Similarity 67.6%; Pred. No. 3.8e-35;  
 Matches 71; Conservative 1; Mismatches 8; Indels 25; Gaps 2;

Qy 1 MKVTAFNPGFNRRVFIKRVSNVTHGRRIDIFASKNFHLQKNTIGTGRKNNRIWLQ 60  
 |||

Db 1 MKVTVAFNQGPNNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRIS----- 53  
 QY 61 FAKLTGFTLMGRRLKMPMYIAGYKTPGRVDDGIIAAYONPASWK 105  
 Db 54 -LKLTSGKIA-----SRKVDGIIAAYONPASWK 80

## RESULT 4

AAB69120  
 ID AAB69120 standard; Protein; 214 AA.

AC AAB69120;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:61.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiSugiPollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.

DR WPI, 2001-185061/19.

DR N-PSDB; AAF59044.

PT Novel peptide and its use -

PS Example 11; Page 58-59; 75PD; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7

CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an antiSugiPollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.

XX Sequence 214 AA;

Query Match 38.2%; Score 213; DB 22; Length 214;  
 Best Local Similarity 41.0%; Pred. No. 1.6e-19;

Matches 55; Conservative 7; Mismatches 8; Indels 64; Gaps 5;

QY 1 MKVTVAFNQGPNNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGR----- 50  
 Db 30 MKVTVAFNQGP-----DIFASKNFHLQKNTIGTGR----- 50

QY 51 -----RMKNKRITL-----QPAKLTGFTLMGRRLKMPMYIAGYKTPDGRV 91  
 Db 71 HVANNNDPSGKTEGNGITTKKAENVEQFAKLTGFTLMGRA-----DPR-- 115

QY 92 DGIIAAYONPASWK 105  
 Db 116 -GIIAAYONPASWK 128

## RESULT 5

AAB69103

ID AAB69103 standard; Protein; 210 AA.

AC AAB69103;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:12.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiSugiPollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP2000327699-A.

PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.

PR 15-MAR-1999; 99JP-0068316.

PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.

DR WPI, 2001-185061/19.

DR N-PSDB; AAF59012.

PT Novel peptide and its use -

PS Claim 7; Page 39-40; 75PD; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:  
 CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7

CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an antiSugiPollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.

XX Sequence 210 AA;

Query Match 34.5%; Score 192.5; DB 22; Length 210;  
 Best Local Similarity 39.6%; Pred. No. 7.5e-17;

Matches 53; Conservative 7; Mismatches 7; Indels 67; Gaps 6;

QY 1 MKVTVAFNQGPNNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGR----- 50  
 Db 29 MKVTVAFNQGP-----FASKNFHLQKNTIGTGR----- 67

QY 51 -----RMKNKRITL-----QPAKLTGFTLMGRRLKMPMYIAGYKTPDGRV 90  
 Db 68 HVANNNDPSGKTEGNGITTKKAENVEQFAKLTGFTLMGRA-----DPR- 113

QY 91 VDGIIAAYONPASWK 104  
 Db 114 -GIIAAYONPASWK 125

## RESULT 6

AAB69104

ID AAB69104 standard; Protein; 96 AA.

AC AAB69104;

DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:13.

KM Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiSugiPollinosis.

[illegible]

CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.  
 CC  
 SQ Sequence 97 AA;

Query Match 29.8%; Score 162; DB 22; Length 97;  
 Best Local Similarity 39.6%; Pred. No. 2.6e-13;  
 Matches 40; Conservative 7; Mismatches 6; Indels 48; Gaps 3;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGR-----50  
 DB 16 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGR-----50  
 ID AAB69117 standard; Protein; 93 AA.

QY 51 -----RWKNNRIWL-----QPAKLTGFTLMGR 72  
 DB 57 VHVANNNDPSGKYEKGNIYTKKEAFNVEQPAKLTGFTLMGR 97  
 ID AAB69102 standard; Protein; 186 AA.

AC AAB69102;  
 AC  
 XX  
 DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:10.  
 XX  
 XX Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiugipollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP2000327699-A.  
 PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.  
 PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.  
 DR N-PSDB; AAF59011.

PT Novel peptide and its use -

PS Claim 14; Page 36-37; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:

CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an antiugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.

XX Sequence 186 AA;

Query Match 26.8%; Score 149.5; DB 22; Length 186;  
 Best Local Similarity 37.1%; Pred. No. 2.8e-11;  
 Matches 43; Conservative 7; Mismatches 23; Indels 43; Gaps 4;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNT-----IGTGR 50  
 DB 29 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNT-----IGTGR 50  
 ID AAB69117 standard; Protein; 93 AA.

QY 51 RWKNNRIWLQPAKLTGFTLMGRRLKMPYIAGYTFDGRVD--GIIAAYGNPASM 104  
 DB 51 Sugi allergen protein Cryj1 derived epitope for T cells.

DB 68 VHVANNNDPSGKYEKGNIYTKK-----EAFNVEQPAKLTGFTLMGR 113  
 ID AAB69117 standard; Protein; 93 AA.

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGR-----50  
 DB 14 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGTGR-----50  
 ID AAB69117 standard; Protein; 93 AA.

QY 51 -----RWKNNRIWL-----QPAKLTGFTLMGR 72  
 DB 57 VHVANNNDPSGKYEKGNIYTKKEAFNVEQPAKLTGFTLMGR 97  
 ID AAB69102 standard; Protein; 186 AA.

AC AAB69102;  
 AC  
 XX  
 DT 23-APR-2001 (first entry)

DE Cedar pollen allergen T cell epitope derived protein SEQ ID NO:29.  
 XX  
 XX Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiugipollinosis.

OS Cryptomeria japonica.  
 OS Synthetic.

PN JP2000327699-A.  
 PD 28-NOV-2000.

PF 15-MAR-2000; 2000JP-0071710.  
 PR 15-MAR-1999; 99JP-0068316.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY) SANKYO CO LTD.

DR WPI; 2001-185061/19.  
 DR N-PSDB; AAF59015.

PT Novel peptide and its use -

PS Example 10; Page 46; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
 CC its polymerizate, where the peptide (I) has a formula of:

CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
 CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
 CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
 CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
 CC allergens. The peptide can be used in an antiugipollinosis agent.  
 CC AAF59011 to AAF59062 and AAB69094 to AAB69121 represent sequences used  
 CC in the exemplification of the present invention.

XX Sequence 93 AA;

Query Match 25.4%; Score 141.5; DB 22; Length 93;  
 Best Local Similarity 37.6%; Pred. No. 1.2e-10;  
 Matches 38; Conservative 7; Mismatches 5; Indels 51; Gaps 4;

QY 1 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNT-----IGTGR 50  
 DB 14 MKVTAFNFGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNT-----IGTGR 50  
 ID AAB69117 standard; Protein; 93 AA.

QY 51 -----RWKNNRIWL-----QPAKLTGFTLMGR 72  
 DB 53 VHVANNNDPSGKYEKGNIYTKKEAFNVEQPAKLTGFTLMGR 93  
 ID AAB69102 standard; Protein; 186 AA.

KW T cell epitope; sugi allergen proteins Cryj1; Cryj2; treatment;  
 KM sugi-pollinosis; allergic reaction; pollen.  
 XX Synthetic.  
 OS  
 XX JP10259198-A.  
 PN  
 XX 29-SEP-1998.  
 PD  
 XX 22-DEC-1997; 97JP-0353448.  
 PF  
 XX 24-DEC-1996; 96JP-0343441.  
 PR  
 XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.  
 XX WPI; 1998-577037/49.  
 DR  
 XX A linked T cell epitope peptide - used for the treatment of  
 PT sugi-pollinosis  
 PS Claim 10; Page 5; 21pp; Japanese.  
 PS  
 XX AAW80339-58 represent epitopes for T cells, derived from the sugi  
 CC allergen proteins Cryj1 (AAW80339-44, AAW80350-53 and AAW80356-58) and  
 CC Cryj2 (AAW80345-49 and AAW80354-55). The peptides are useful for the  
 CC treatment of sugi-pollinosis, an allergic reaction of the body to  
 CC pollen.  
 CC  
 XX Sequence 47 AA;  
 SQ  
 Query Match 23.6%; Score 131.5; DB 19; Length 47;  
 Best Local Similarity 33.7%; Pred. No. 1e-09;  
 Matches 35; Conservative 0; Mismatches 0; Indels 69; Gaps 2;  
 QY 1 MKVTAFNFGFNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGGRKNNRIWLQ 60  
 DB 13 MKVTAFNFGF-----FASKNFHLQKNT----- 35  
 QY 61 FKLNGFTLMGRRLKMPMYIAGYKTFPGRVDGIITAAVONPASM 104  
 DB 36 -----GIIAAYONPASM 47  
 RESULT 12  
 AAB23905  
 ID AAB23905 standard; peptide; 74 AA.  
 AC AAB23905;  
 XX  
 DT 17-JAN-2001 (first entry)  
 DE Artificial sequence designed peptide #26.  
 KM Peptide synthesis; chemical synthesis; solid phase synthesis.  
 XX Synthetic.  
 OS  
 XX WO200055182-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000WO-JP01584.  
 PF  
 XX 15-MAR-1999; 99JP-0067917.  
 PR  
 XX (SANY ) SANKYO CO LTD.  
 PA Kawaguchi J, Serizawa N;  
 PI WPI; 2000-602106/57.  
 DR  
 XX Chemically synthesizing a peptide by solid phase synthesis by adding

PT protective group to inactivate part of alpha amino group of an amino  
 PT acid group -  
 XX  
 XX Example 6; Page 23; 38pp; Japanese.  
 XX  
 CC The present invention describes a method for chemically synthesizing a  
 CC peptide by solid phase synthesis. The method comprises processing a  
 CC carrier resin to which the C-terminal amino acid derivative of the  
 CC target peptide has been preliminarily bonded is characterised by adding  
 CC a protective group to inactivate a part of alpha-amino groups of an  
 CC amino acid capable of undergoing a peptide extension reaction on the  
 CC carrier resin. The method is useful for synthesizing a peptide by solid  
 CC phase synthesis. Long chain peptides can be readily synthesised in with  
 CC a high success ratio without using the segment condensation method or  
 CC gene manipulations. The present sequence represents an amino acid  
 CC peptide sequence given in an example from the present invention.  
 SQ  
 Sequence 74 AA;  
 Query Match 21.6%; Score 120.5; DB 21; Length 74;  
 Best Local Similarity 52.0%; Pred. No. 5.1e-08;  
 Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;  
 QY 1 MKVTAFNFGFNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNTIGGR 50  
 DB 6 MKVTAFNFGF-----DIFASKNFHLQKNTIGGR 36  
 RESULT 13  
 AAB23906  
 ID AAB23906 standard; peptide; 94 AA.  
 AC AAB23906;  
 XX  
 DT 17-JAN-2001 (first entry)  
 DE Artificial sequence designed peptide #27.  
 KM Peptide synthesis; chemical synthesis; solid phase synthesis.  
 XX Synthetic.  
 OS  
 XX WO200055182-A1.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000WO-JP01584.  
 PF  
 XX 15-MAR-1999; 99JP-0067917.  
 PR  
 XX (SANY ) SANKYO CO LTD.  
 PA Kawaguchi J, Serizawa N;  
 PI WPI; 2000-602106/57.  
 DR  
 XX Chemically synthesizing a peptide by solid phase synthesis by adding  
 PT protective group to inactivate part of alpha amino group of an amino  
 PT acid group -  
 XX  
 XX Example 6; Page 24; 38pp; Japanese.  
 XX  
 CC The present invention describes a method for chemically synthesizing a  
 CC peptide by solid phase synthesis. The method comprises processing a  
 CC carrier resin to which the C-terminal amino acid derivative of the  
 CC target peptide has been preliminarily bonded is characterised by adding  
 CC a protective group to inactivate a part of alpha-amino groups of an  
 CC amino acid capable of undergoing a peptide extension reaction on the  
 CC carrier resin. The method is useful for synthesizing a peptide by solid  
 CC phase synthesis. Long chain peptides can be readily synthesised in with  
 CC a high success ratio without using the segment condensation method or  
 CC gene manipulations. The present sequence represents an amino acid  
 CC peptide sequence given in an example from the present invention.

XX Sequence 94 AA;  
SQ

Query Match 21.6%; Score 120.5; DB 21; Length 94;  
Best Local Similarity 52.0%; Pred. No. 7.2e-08;  
Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;

QY 1 MKVTVAENQFGPNRRVFIKRVSNVHIGRRIDIFASKNFHLQKNLTGTGR 50  
DB 26 MKVTVAENQFGP-----DIFASKNFHLQKNLTGTGR 56

## RESULT 14

AAB23897  
ID AAB23897 standard; peptide; 95 AA.

AC AAB23897;

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide SEQ ID NO:5.

KM Peptide synthesis; chemical synthesis; solid phase synthesis.

OS Synthetic.

PN WO200055182-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-JP01584.

PR 15-MAR-1999; 99JP-0067917.

PA (SANY) SANKYO CO LTD.

PI Kawaguchi J, Serizawa N;

DR WPI; 2000-602106/57.

PT Chemically synthesizing a peptide by solid phase synthesis by adding  
PT protective group to inactivate part of alpha amino group of an amino  
PT acid group

PS Example 5; Page 34-35; 38pp; Japanese.

CC The present invention describes a method for chemically synthesizing a  
CC peptide by solid phase synthesis. The method comprises processing a  
CC carrier resin to which the C-terminal amino acid derivative of the  
CC target peptide has been preliminarily bonded is characterized by adding  
CC a protective group to inactivate a part of alpha-amino groups of an  
CC amino acid capable of undergoing a peptide extension reaction on the  
CC carrier resin. The method is useful for synthesizing a peptide by solid  
CC phase synthesis. Long chain peptides can be readily synthesised in with  
CC a high success ratio without using the segment condensation method or  
CC gene manipulations. The present sequence represents an amino acid  
CC peptide sequence given in an example from the present invention.

SQ Sequence 95 AA;

Query Match 21.6%; Score 120.5; DB 21; Length 95;  
Best Local Similarity 52.0%; Pred. No. 7.2e-08;  
Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;

QY 1 MKVTVAENQFGPNRRVFIKRVSNVHIGRRIDIFASKNFHLQKNLTGTGR 50  
DB 27 MKVTVAENQFGP-----DIFASKNFHLQKNLTGTGR 57

RESULT 15  
AAB23901  
ID AAB23901 standard; peptide; 95 AA.  
XX

AC AAB23901;

DT 17-JAN-2001 (first entry)

DE Artificial sequence designed peptide #23.

KM Peptide synthesis; chemical synthesis; solid phase synthesis.

OS Synthetic.

PN WO200055182-A1.

PD 21-SEP-2000.

PF 15-MAR-2000; 2000WO-JP01584.

PR 15-MAR-1999; 99JP-0067917.

PA (SANY) SANKYO CO LTD.

PI Kawaguchi J, Serizawa N;

DR WPI; 2000-602106/57.

PT Chemically synthesizing a peptide by solid phase synthesis by adding  
PT protective group to inactivate part of alpha amino group of an amino  
PT acid group

PS Example 5; Page 21-22; 38pp; Japanese.

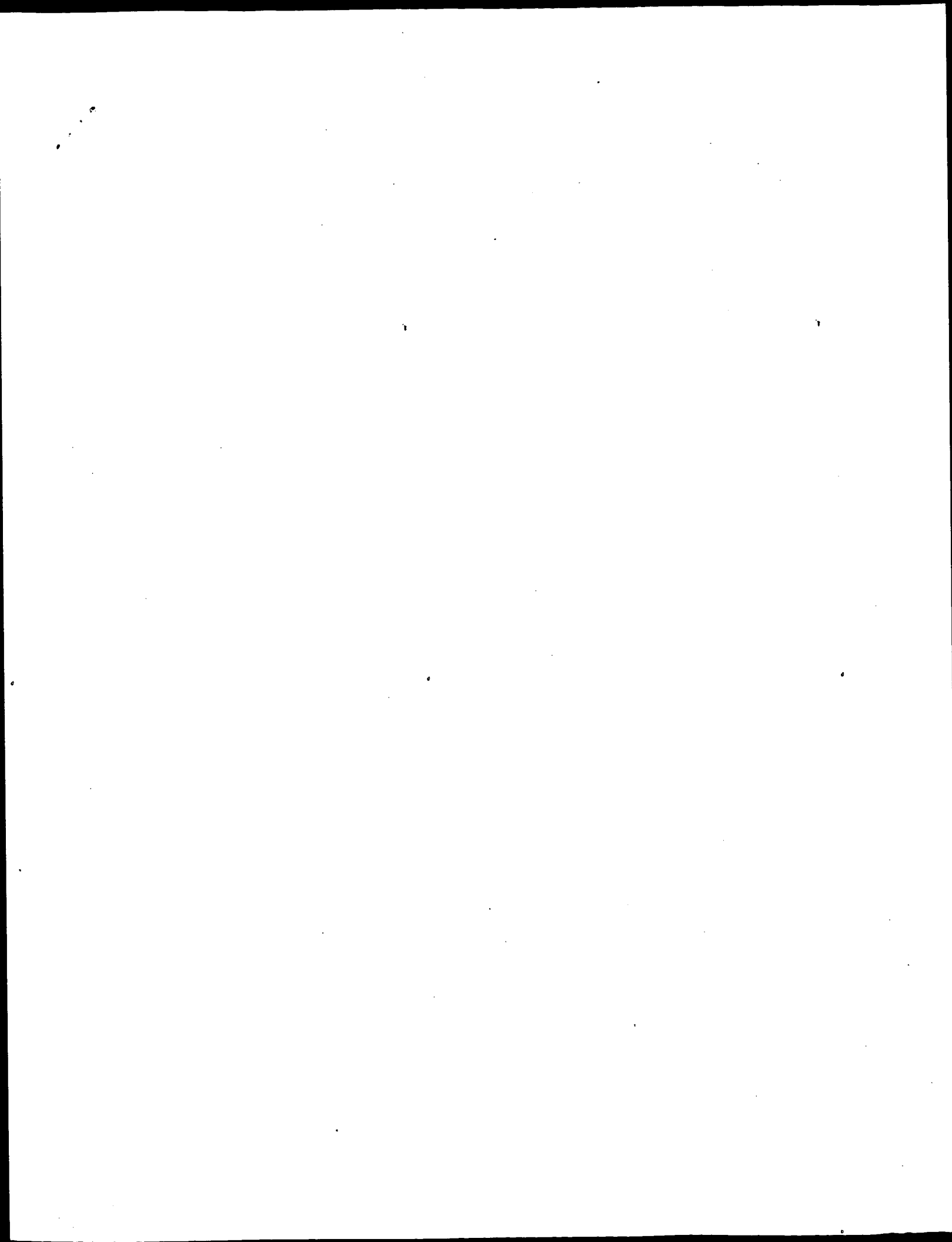
CC The present invention describes a method for chemically synthesizing a  
CC peptide by solid phase synthesis. The method comprises processing a  
CC carrier resin to which the C-terminal amino acid derivative of the  
CC target peptide has been preliminarily bonded is characterized by adding  
CC a protective group to inactivate a part of alpha-amino groups of an  
CC amino acid capable of undergoing a peptide extension reaction on the  
CC carrier resin. The method is useful for synthesizing a peptide by solid  
CC phase synthesis. Long chain peptides can be readily synthesised in with  
CC a high success ratio without using the segment condensation method or  
CC gene manipulations. The present sequence represents an amino acid  
CC peptide sequence given in an example from the present invention.

SQ Sequence 95 AA;

Query Match 21.6%; Score 120.5; DB 21; Length 95;  
Best Local Similarity 52.0%; Pred. No. 7.2e-08;  
Matches 26; Conservative 3; Mismatches 2; Indels 19; Gaps 1;

QY 1 MKVTVAENQFGPNRRVFIKRVSNVHIGRRIDIFASKNFHLQKNLTGTGR 50  
DB 27 MKVTVAENQFGP-----DIFASKNFHLQKNLTGTGR 57

Search completed: April 20, 2003, 13:06:03  
Job time : 128.382 secs





GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 17 Seconds  
(without alignments)  
757.765 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKVTAFNQFGNRRVFIKR.....KFIRRVGIIAAYGNPASWK 134

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	188	26.4	374	2 JC2124	major allergen Cry
2	188	26.4	374	2 JC2123	major allergen Cry
3	148.5	20.9	514	2 S48730	Cry j II protein -
4	148.5	20.9	514	2 JC2498	second major aller
5	118.5	16.7	514	2 JC7100	polygalacturonase
6	110.5	15.5	542	2 TC6728	pectate lyase (EC
7	107.5	15.1	507	2 JC7366	Jun a 2 protein -
8	104.5	14.7	463	2 T46165	pectate lyase-like
9	103	14.5	404	2 S12209	pectate lyase (EC
10	99	13.5	472	2 T51456	pectate lyase-like
11	98.5	13.9	396	2 A39099	allergen Amb a 1.1
12	97	13.6	397	2 C53240	allergen Amb a 1.3
13	97	13.6	397	2 C39099	allergen Amb a 1.3
14	97	13.6	455	2 TC00856	hypothetical prote
15	95.5	13.4	459	2 T05556	pectate lyase (EC
16	94.5	13.3	404	2 T05556	hypothetical prote
17	94.5	13.3	431	2 F86179	pectate lyase (EC
18	93	13.1	397	2 S26211	hypothetical prote
19	93	13.1	398	2 B39099	pectate lyase (EC
20	93	13.1	450	2 TC09524	allergen Amb a 1.2
21	92	12.9	397	2 E53240	probable pectate 1
22	92	12.9	398	2 B53240	allergen Amb a 1.1
23	91.5	12.9	368	2 G86427	probable pectate 1
24	91	12.8	398	2 T07058	pectate lyase (EC
25	89	12.5	392	2 D53240	allergen Amb a 1.4
26	89	12.5	434	2 S29612	pectate lyase (EC
27	88	12.4	438	2 S4335	pectate lyase (EC
28	84	11.8	374	2 T05240	pectate lyase (EC
29	84	11.8	374	2 H85148	probable pectate 1

30	84	11.8	390	2 H86253	hypothetical prote
31	84	11.8	394	2 T49115	pectate lyase like
32	82.5	11.6	449	2 S27098	pectate lyase (EC
33	82	11.5	394	2 T49116	pectate lyase like
34	81	11.4	219	2 C84647	probable synaptobr
35	79	11.1	274	2 T03932	pectate lyase (EC
36	78.5	11.0	418	2 T07701	pectate lyase (EC
37	78	11.0	341	2 T47653	pectate lyase-like
38	76.5	10.8	648	2 S10869	enterotoxin A - Cl
39	76	10.7	398	2 F72335	hypothetical prote
40	74.5	10.5	368	2 F83785	required for glyco
41	74	10.4	219	2 T04630	synaptobrevin homo
42	71.5	10.1	657	2 T22451	hypothetical prote
43	71	10.0	2710	2 A37052	coxin A - Clostrid
44	70.5	9.9	622	2 S17402	paraspinal crystal
45	70.5	9.9	1003	2 T26746	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC2124  
major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar  
C:Species: Cryptomeria japonica (Japanese cedar)  
C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000  
C:Accession: JC2124  
R:Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.  
Biochem. Biophys. Res. Commun. 199, 619-625, 1994  
A>Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese  
A:Reference number: JC2123; NCID:94183234; PMID:8135802  
A:Accession: JC2124  
A:Molecule type: mRNA  
A:Residues: 1-374 <SON>  
A:Cross-references: GB:D26545; NID:G493633; PIDN:BA0543.1; PID:G493634  
A:Experimental source: pollen  
A>Note: the authors described carbohydrate binding site for residue 279  
C:Superfamily: pectate lyase LATS9  
C:Keywords: glycoprotein; pollen  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>  
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%; Score 188; DB 2; Length 374;  
Best Local Similarity 71.2%; Pred. No. 1.9e-12;  
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 73 RPLWIFSGNMNITKMKPMYIAGYKTFDGRRAEVSYVHVGAKFIRRVGII 124  
DB 87 RPLWIFSGNMNITKMKPMYIAGYKTFDGRRAEVSYVHVGAKFIRRVGII 138

RESULT 2  
JC2123  
major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar  
C:Species: Cryptomeria japonica (Japanese cedar)  
C>Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000  
C:Accession: JC2123; PC2065  
R:Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.  
Biochem. Biophys. Res. Commun. 199, 619-625, 1994  
A>Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese  
A:Reference number: JC2123; NCID:94183234; PMID:8135802  
A:Accession: JC2123  
A:Molecule type: mRNA  
A:Residues: 1-374 <SON>  
A:Cross-references: GB:D26544; NID:G493631; PIDN:BA0543.1; PID:G493632  
A:Experimental source: pollen  
A:Accession: PC2065  
A:Molecule type: protein  
A:Residues: 22-53;58-81;219-232;236-258;299-307;346-372 <S02>  
A>Note: the authors described carbohydrate binding site for residue 279  
C:Superfamily: pectate lyase LATS9  
C:Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-3/4/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>  
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.4%; Score 188; DB 2; Length 374;  
Best Local Similarity 71.2%; Pred. No. 1.9e-12;  
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 73 RPLMIIFSGNNIKLKMPMYIAGYKTFDGRRAEVSYYHVGAKFIRRVDCII 124  
DB 87 RPLMIIFSGNNIKLKMPMYIAGYKTFDGRGAQYVIGNGPCVFIRKVSNI 138

## RESULT 3

S48730  
Cry j II protein - Japanese cedar  
C:Species: Cryptomeria japonica (Japanese cedar)  
C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999  
C:Accession: S48730  
R:Namba, M.; Kurose, M.; Torigoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kur  
FEBB Lett. 353, 124-128, 1994  
A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar  
A:Reference number: S48730; MUID:95010777; PMID:7926035  
A:Accession: S48730  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-514 <NM>  
A:Cross-references: GB:D37765; NID:9577695; PIDN:BA07021.1; PID:d1007598; PID:g577696

Query Match 20.9%; Score 148.5; DB 2; Length 514;  
Best Local Similarity 41.4%; Pred. No. 4.9e-08;  
Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLMIFSGNNIKLKMP 90  
DB 236 IDIFASKNFHLQKNTIGT-----DDCVAIIGTSSNIYIEDL 272

QY 91 MYIAGYKTFD-----RRAEVSYYHVGAKFIRRVDCI 123  
DB 273 ICGPHGISISLGRNSRAEVSYYHVGAKFIDTONGL 311

## RESULT 4

JC2498  
second major allergen Cry j II precursor - Japanese cedar  
C:Species: Cryptomeria japonica (Japanese cedar)  
C:Date: 16-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 21-Jul-2000  
C:Accession: JC2498; PC2346; A60147

R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.  
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994  
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese  
A:Reference number: JC2498; MUID:94271186; PMID:8002972  
A:Accession: JC2498  
A:Molecule type: mRNA  
A:Residues: 1-514 <XM>  
A:Cross-references: DDBJ:DB29772; NID:g506857; PIDN:BA06172.1; PID:g506858

A:Accession: PC2346  
A:Molecule type: protein  
A:Residues: 52-61 <KO2>  
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhashi, T.  
Allergy 45, 309-312, 1990  
A:Title: Identification of the second major allergen of Japanese cedar pollen.  
A:Reference number: A60147; MUID:90342988; PMID:2362797  
A:Accession: A60147

A:Molecule type: protein  
A:Residues: 55-64 <SAK>  
C:Keywords: glycoprotein; pollen  
F:1-54/Domain: signal sequence #status predicted <SIG>  
F:55-46/Product: second major allergen Cry j #status predicted <MAT>  
F:428,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.9%; Score 148.5; DB 2; Length 514;  
Best Local Similarity 41.4%; Pred. No. 4.9e-08;

Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLMIFSGNNIKLKMP 90  
DB 236 IDIFASKNFHLQKNTIGT-----DDCVAIIGTSSNIYIEDL 272

QY 91 MYIAGYKTFD-----RRAEVSYYHVGAKFIRRVDCI 123  
DB 273 ICGPHGISISLGRNSRAEVSYYHVGAKFIDTONGL 311

## RESULT 5

JC7100  
polygalacturonase Cha o 2 - Japanese cypress  
C:Species: Chamaecyparis obtusa (Japanese cypress)  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-May-2000  
C:Accession: JC7100; PC7026  
R:Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.  
Biochem. Biophys. Res. Commun. 263, 166-171, 1999  
A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all  
A:Reference number: JC7100; MUID:99417540; PMID:10486272  
A:Accession: JC7100  
A:Molecule type: mRNA  
A:Residues: 1-514 <MO>  
A:Accession: PC7026  
A:Molecule type: protein  
A:Residues: 51-62 <MO2>

Query Match 16.7%; Score 118.5; DB 2; Length 514;  
Best Local Similarity 35.4%; Pred. No. 8.3e-05;  
Matches 35; Conservative 10; Mismatches 25; Indels 29; Gaps 2;

QY 31 IDIFASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLMIFSGNNIKLKMP 90  
DB 236 IDIFASKNFHLQKNTIGT-----DDCVAIIGTSSNIYIEDL 272

QY 91 MYIAGYKTFD-----DGRRAEVSYYHVGAKFIRRVDCI 123  
DB 273 ICGPHGISISLGRNSRAEVSYYHVGAKFIDTONGL 311

## RESULT 6

T06728  
pectate lyase (EC 4.2.2.2) F28P10.100 - Arabidopsis thaliana  
A:Alternate names: protein F28P10.100  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999  
C:Accession: T06728  
R:Quellier, F.; Cholme, N.; Robert, C.; Broctier, P.; Wincker, P.; Cattivico, L.; Artigu  
submitted to the Protein Sequence Database, April 1999  
A:Reference number: Z15793  
A:Accession: T06728  
A:Molecule type: DNA  
A:Residues: 1-542 <OUE>  
A:Cross-references: EMBL:AL049655  
A:Experimental source: cultivar Columbia; BAC clone F28P10

C:Genetics:  
A:Map position: 3  
A:introns: 46/2; 346/3; 413/2; 480/2  
A:Note: F28P10.100  
C:Superfamily: pectate lyase LAR59  
C:Keywords: carbon-oxygen lyase

Query Match 15.5%; Score 110.5; DB 2; Length 542;  
Best Local Similarity 43.3%; Pred. No. 0.00064;  
Matches 26; Conservative 9; Mismatches 20; Indels 5; Gaps 2;

QY 67 FTLGRRLMIFSGNNIKLKMPYIAGYKTFDGRRAEVSYYHVGAKFIRRVDCII 124  
DB 139 YGVIOEPLMIVFSSNMLIRKQELIINSYTLDRG---SAVHITNGCLTLQYVHHI 195

## RESULT 7

UC7366  
 Jun a 2 protein - mountain cedar  
 C:Species: Juniperus ashei (mountain cedar)  
 C:Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 17-Nov-2000  
 C:Accession: UC7366, PC7093  
 R:Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.  
 Biochem. Biophys. Res. Commun. 275, 195-202, 2000  
 A:Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all  
 A:Reference number: UC7366  
 A:Accession: UC7366  
 A:Molecule type: mRNA  
 A:Residues: 1-507 <YOK>  
 A:Cross-references: GB:AJ404653  
 A:Accession: PC7093  
 A:Molecule type: protein  
 A:Residues: 55-63 <Y02>  
 C:Comment: This protein, a second major allergen of mountain cedar pollen, which is inv  
 C:Keywords: glycoprotein; pollen

Query Match 15.1%; Score 107.5; DB 2; Length 507;  
 Best Local Similarity 32.3%; Pred. No. 0.0012;  
 Matches 32; Conservative 13; Mismatches 25; Indels 29; Gaps 2;

QY 31 IDIASKNFHLQKNTIGTGRKKNRIMLOPAKLTGFTLMGRPLWIFSGNNMILKMP 90  
 DB 237 IDIASKRFETKCTIGT-----DDCVAVGTSSNITIKDL 273

QY 91 MYIAGYKTFDG-----RRAEVSYVHNGAKETRRVDI 123  
 DB 274 TCGPGHGMSTIGSLCKGSRSEVSVFVHDAKFIIDTONGL 312

RESULT 8  
 T46165  
 pectate lyase-like protein - Arabidopsis thaliana  
 N:Alternate names: protein T4p2.120  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 02-Sep-2000  
 C:Accession: T46165  
 R:Yakutara, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;  
 submitted to the Protein Sequence Database, December 1999  
 A:Reference number: Z23025  
 A:Accession: T46165  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-463 <NYA>  
 A:Cross-references: EMBL:AL132958  
 A:Experimental source: cultivar Columbia; BAC clone T4D2  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 27/2; 112/2; 182/1; 265/3; 329/3; 399/2  
 A:Note: T4D2.120  
 C:Superfamily: pectate lyase LAT59

Query Match 14.7%; Score 104.5; DB 2; Length 463;  
 Best Local Similarity 30.5%; Pred. No. 0.0024;  
 Matches 32; Conservative 8; Mismatches 32; Indels 33; Gaps 3;

QY 52 WKNNRIML-----QFAKLGTG-----FTLMGRPLMIIF 79  
 DB 66 WPNNRQGLADCGIGFGVALGCKGQFYFVTDSDDAVNPGRITRYGVIOEPLMIIF 125  
 QY 80 SGNNMILKMPMYIAGYKTFDGRRAEVSYVHNGAKETRRVDGII 124  
 DB 126 PSNNMILKQELIENSYKTLDRGANVHIVG-CGCTITQYVANI 169

RESULT 9  
 S12209  
 pectate lyase (EC 4.2.2.2) - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 23-Jul-1999

C:Accession: S12209  
 R:Budelier, K.A.; Smith, A.G.; Gasser, C.S.  
 Mol. Gen. Genet. 224, 183-192, 1990  
 A:Title: Regulation of a stylar transmitting tissue-specific gene in wild-type and trans  
 A:Reference number: S12209; MUID:91117185; PMID:2277637  
 A:Accession: S12209  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-404 <BUD>  
 A:Cross-references: GB:X55193; NID:g19161; PID:CA38979.1; PID:g19162  
 C:Superfamily: pectate lyase LAT59  
 C:Keywords: carbon-oxygen lyase

Query Match 14.5%; Score 103; DB 2; Length 404;  
 Best Local Similarity 30.8%; Pred. No. 0.0029;  
 Matches 33; Conservative 13; Mismatches 31; Indels 30; Gaps 4;

QY 28 GRRIIDPASKNFHLQKNTIGTGRKKNRIM-----LOPAKLTGFTLM 70  
 DB 67 GKSPAFSTICAIQFGKNAIG---GKNRIVVTDSGNDPVPKPTLRHA-----VI 116

QY 71 GRPLMIIFSGNNMILKMPMYIAGYKTFDGRRAEVSYVHNGAKFI 117  
 DB 117 QDEPLMIIFKRDVMIQLKQELVMSYKTLDRGAS---VHISGPCI 160

RESULT 10  
 T51456  
 pectate lyase-like protein - Arabidopsis thaliana  
 N:Alternate names: protein F2G14.230  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 17-Nov-2000  
 C:Accession: T51456  
 R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; New  
 submitted to the Protein Sequence Database, August 2000  
 A:Reference number: Z25394  
 A:Accession: T51456  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-472 <SAT>  
 A:Cross-references: EMBL:AL391146  
 A:Experimental source: cultivar Columbia; BAC clone F2G14  
 C:Genetics:  
 A:Map position: 5  
 A:Introns: 65/2; 316/3; 398/3  
 A:Note: F2G14\_230  
 C:Superfamily: pectate lyase LAT59

Query Match 13.9%; Score 99; DB 2; Length 472;  
 Best Local Similarity 28.2%; Pred. No. 0.0094;  
 Matches 29; Conservative 17; Mismatches 41; Indels 16; Gaps 3;

QY 34 FASKNFHLQKNTIGTGR--WKNNRIMLOPAKLT-----TGFTLMGRPLMIIF 79  
 DB 134 WAKRRKLTLCVSGFGHRTTGKRGRIYVTSNLDMDVNPGRITRHAIVQKEPLMIIF 193  
 QY 80 SGNNMILKMPMYIAGYKTFDGRRAEVSYVHNG--AKTIRRV 120  
 DB 194 KNDMSIRLNQELLINSHKITDRGANVHVAHGATMGQVKNV 226

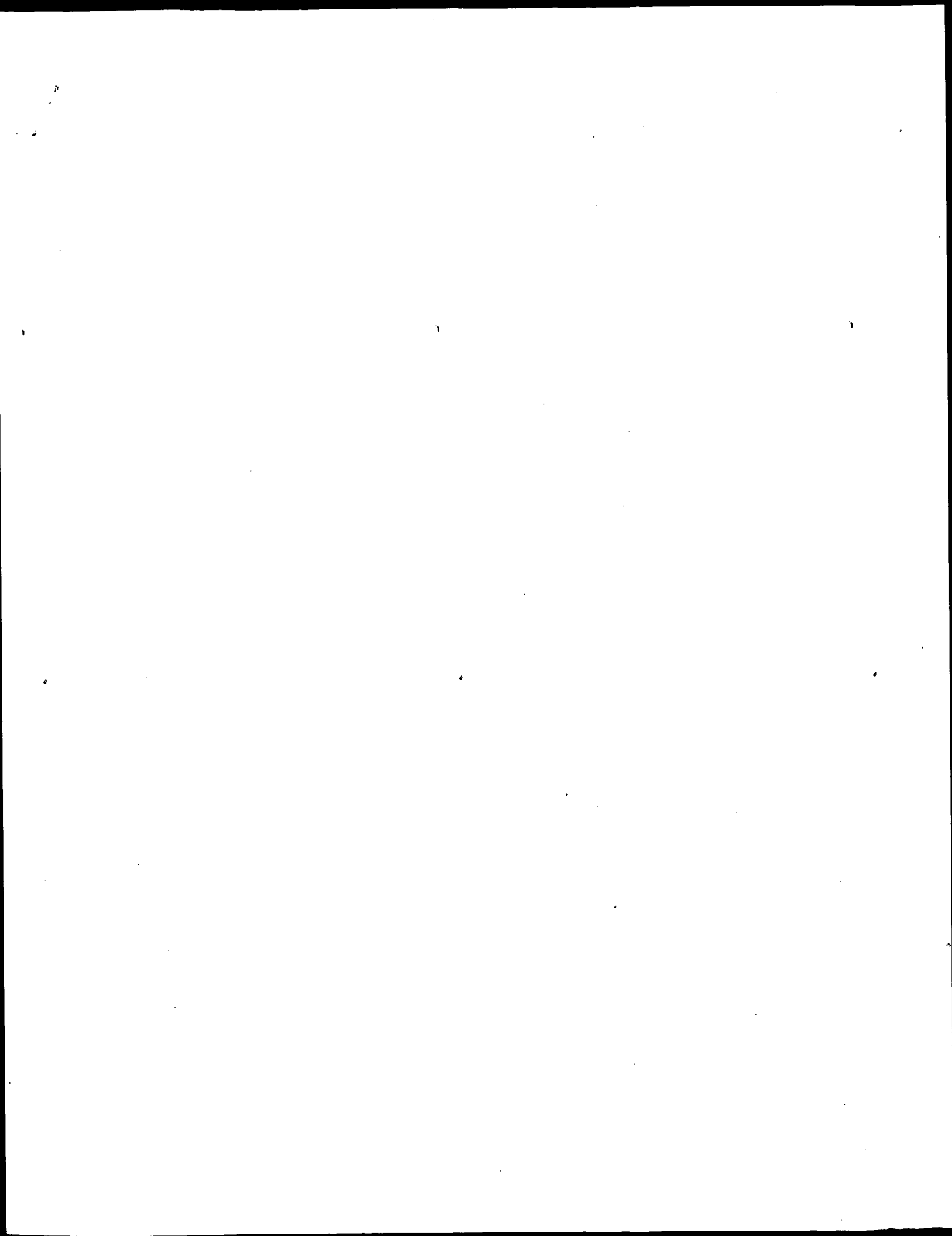
RESULT 11  
 A39099  
 allergen Amb a 1 precursor - common ragweed  
 C:Species: Ambrosia artemisiifolia (common ragweed)  
 C:Date: 27-Nov-1991 #sequence\_revision 03-Apr-1992 #text\_change 29-Sep-1999  
 C:Accession: A39099; A60895; A53240  
 R:Rafnar, T.; Grifflith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.  
 J. Biol. Chem. 266, 1229-1236, 1991  
 A:Title: Cloning of Amb a 1 (antigen E), the major allergen family of short ragweed pol  
 A:Reference number: A39099; MUID:91093235; PMID:1702434  
 A:Accession: A39099  
 A:Molecule type: mRNA



C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 01-Mar-2002  
 C/Accession: G86278  
 R/Theologian, A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Huizal, L.  
 Nature 408, 816-820, 2000  
 A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maitl, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Sakano, H.; Sakano, H.; Sakano, H.; Sakano,  
 A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/Reference number: A86141; MUID:21016719; PMID:11130712  
 A/Accession: G86278  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-459 <STO>  
 A/Cross-reference: GB:AE005172; NID:97262684; PIDN:AAF43942.1; GSPDB:GN00141  
 C/Genetics:  
 A/Map position: 1  
 C/Superfamily: pectate lyase LAT59

Query Match 13.4%; Score 95.5; DB 2; Length 459;  
 Best Local Similarity 28.6%; Pred. No. 0.022;  
 Matches 32; Conservative 9; Mismatches 28; Indels 43; Gaps 4;  
 QY 50 RRMKNRIMLQFAKLTGFTL-MGR----- 73  
 DB 115 RNMANNR-----KLLADCVLGFGRRTTGKDGPIYVVKASDNDLINPKPGLRHAVTRD 169  
 QY 74 -PLWIIISGNMNIKLKMPYIAGYTFDGRRAEVSYYVNGAKFIRRVGII 124  
 DB 170 GPLWIIIFARSWIITLQGLMITSDKITDGRGARV-YMEGAGLTLPVNNVI 220

Search completed: April 23, 2003, 18:08:13  
 Job time : 19 secs



GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 13 Seconds

(without alignments)  
427.525 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKVTAFNQGPNRRVFIKR.....KPIRRVGIILAAQNPASWK 134

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	26.4	374	1	SBP_CRYJA
2	150	21.1	367	1	MPA1_JUNPERUS
3	148.5	20.9	514	1	MPA2_CRYJA
4	143	20.1	346	1	MPA1_CUPAR
5	142	20.0	375	1	MPA1_CHAOS
6	103	14.5	404	1	MP12_LYCES
7	98.5	13.9	396	1	MP11_AMBAR
8	97	13.6	397	1	MP13_AMBAR
9	93	13.1	397	1	PEL_TOBAC
10	93	13.1	398	1	MP12_AMBAR
11	92	12.9	397	1	MPA2_AMBAR
12	91	12.8	398	1	PESE_LYCES
13	89	12.5	392	1	MP14_AMBAR
14	89	12.5	434	1	PEL_LILLO
15	82.5	11.6	449	1	PESE_LYCES
16	81.5	11.5	234	1	A298_DRONE
17	80.5	11.3	605	1	PHAC_METEX
18	71	10.0	2710	1	TOXA_CLODI
19	70.5	9.9	622	1	C2AC_BACTU
20	69.5	9.8	338	1	YUX8_YEAST
21	68.5	9.6	234	1	A298_DRONE
22	68.5	9.6	551	1	CHIT_NPVAC
23	67.5	9.5	985	1	AGLU_ASFOR
24	67.5	9.5	1070	1	YHVA_YEAST
25	67	9.4	757	1	YHVA_YEAST
26	66.5	9.4	110	1	Y12K_SMSV4
27	66	9.3	375	1	Y4ED_RHISM
28	66	9.3	382	1	PHYT_BACST
29	66	9.3	382	1	NODO_RHIS3
30	65	9.1	576	1	DPEP_SOLTU
31	65	9.1	726	1	YB1D_SCHPO
32	65	9.1	781	1	TL22_CHICK
33	64.5	9.1	1010	1	CLPP_CHLEU

34	64	9.0	532	1	PUR2_ARATH	P52420 arabidopsis
35	64	9.0	633	1	NODO_RHISB	O07309 r nodd bifu
36	63.5	8.9	196	1	YMO7_YEAST	Q04487 saccharomyc
37	63.5	8.9	309	1	Y4OQ_RHISM	P55602 rhizobium s
38	63.5	8.9	326	1	SCER_PEDPE	P43472 pedicoccus
39	63	8.9	347	1	NI2M_RABIT	O79428 oryctolagus
40	63	8.9	477	1	GIGA_SYNT3	P74521 synecocyst
41	63	8.9	522	1	CPV1_ORENI	P70091 oreochromis
42	63	8.9	725	1	GUNG_CLOCE	P37700 clostridium
43	63	8.9	755	1	COMP_RAT	P35444 rattus norv
44	63	8.9	1121	1	MDRE_HUMAN	Ognms homo sapien
45	62.5	8.8	262	1	FLGG_AGRIS	Q44338 agrobacteri

## ALIGNMENTS

RESULT 1  
ID SBP\_CRYJA STANDARD; PRT; 374 AA.  
AC P18632;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Sugt basic protein precursor (SBP) (Major allergen Cry j 1) (Cry j 1).  
OS Cryptomeria japonica (Japanese cedar).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.  
OK NCBI\_TaxID=3369;  
[1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Pollen;  
RX MEDLINE=94183234; PubMed=8135802;  
RA Kono K.; Komiyama N.; Shimizu K.; Kusakabe T.; Morikubo K.;  
RT "Cloning and sequencing of cDNA coding for Cry j 1, a major allergen  
of Japanese cedar pollen.";  
RL Biochem. Biophys. Res. Commun. 199:619-625(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pollen;  
RX Namba M.; Kurose M.; Torigoe K.; Fukuda S.; Kurimoto M.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 22-41.  
RC TISSUE=Pollen;  
RX MEDLINE=89031257; PubMed=3181436;  
RA Tanai M.; Ando S.; Usui M.; Kurimoto M.; Sakaguchi M.; Inoue S.;  
RT "N-terminal amino acid sequence of a major allergen of Japanese cedar  
pollen (Cry j 1).";  
RL FEBS Lett. 239:329-332(1988).  
RN [4]  
RP CARBOHYDRATES.  
RC TISSUE=Pollen;  
RX MEDLINE=95003748; PubMed=7920021;  
RA Hijikata A.; Matsumoto I.; Kojima K.; Ogawa H.;  
RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar  
(Cryptomeria japonica) pollen allergen, Cry j 1.";  
RL Int. Arch. Allergy Immunol. 105:198-202(1994).  
RN [5]  
RP STRUCTURE OF CARBOHYDRATES.  
RC TISSUE=Pollen;  
RX MEDLINE=95332249; PubMed=7608114;  
RA Hino K.; Yamamoto S.; Sano O.; Taniguchi Y.; Kohno K.; Usui M.;  
RT "Carbohydrate structures of the glycoprotein allergen Cry j 1 from  
Japanese cedar (Cryptomeria japonica) pollen.";  
RL J. Biochem. 117:289-295(1995).  
CC -1- PTM: CONTAINS FUROSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.  
CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR  
POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.  
CC -1- MISCELLANEOUS: THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM

B DIFFERS IN SIX POSITIONS.  
 -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 AMB A I/AMB A II/CRY J I SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL, D26544; BAA05542.1; -  
 DR EMBL, D26545; BAA05543.1; -  
 DR EMBL, D34639; BAA07020.1; -  
 DR PIR, A44773; A44773.  
 DR GlycoSuiteDB; P18632; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec lyase; 1.  
 DR PRINTS; PR00807; AMBALLERGEN.  
 DR Allergen; Glycoprotein; Multigene family; Signal.  
 KW SIGNAL  
 FT CHAIN 1 21  
 FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (COMPLEX).  
 FT CARBOHYD 293 293 /FTIG-CAR 000135.  
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (COMPLEX).  
 FT VARIANT 12 12 L -> F (IN CRY J 1-B).  
 FT VARIANT 143 143 H -> Y (IN CRY J 1-B).  
 FT VARIANT 202 202 S -> T (IN CRY J 1-B).  
 FT VARIANT 221 221 L -> S (IN CRY J 1-B).  
 FT VARIANT 358 358 Q -> H (IN CRY J 1-B).  
 FT VARIANT 361 361 K -> Q (IN CRY J 1-B).  
 SQ SEQUENCE 374 AA; 40645 MW; 74AB25950248F56F CRC64;

Query Match 26.4%; Score 188; DB 1; Length 374;  
 Best Local Similarity 71.2%; Pred. No. 2.7e-12;  
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

-----  
 QY 73 RPLMIIFSGMMNLIKMPMTIAGYKTFDGRRAEVSYYVHNGAKFIRVDGII 124  
 DB 87 RPLMIIFSGMMNLIKMPMTIAGYKTFDGRRAEVSYYVHNGAKFIRVDGII 138

RESULT 2  
 MPAL JUNAS STANDARD; PRT; 367 AA.  
 AC P81254; O9ZNU7;  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Major pollen allergen Jun a 1 precursor.  
 OS Juniperus ashei (Osark white cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 OC NCBI\_TaxID=13101;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;  
 RP 117-124; 134-140; 160-164; 256-263 AND 322-325.  
 RC TISSUE=Pollen;  
 RX MEDLINE=99414163; PubMed=10482836;  
 RA Mldoro-Horintu T.M., Goldblum R.M., Kurosky A., Wood T.G.,  
 RA Brooks E.G.;  
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major  
 RT allergen, Jun a 1."  
 RL J. Allergy Clin. Immunol. 104:613-617(1999).  
 RN [2]  
 RP SEQUENCE OF 22-50.  
 RP TISSUE=Pollen;  
 RX MEDLINE=99414162; PubMed=10482835;  
 RA Mldoro-Horintu T., Goldblum R.M., Kurosky A., Goetz D.W.,

RA Brooks E.G.;  
 RT "Isolation and characterization of the mountain cedar (Juniperus  
 RT ashei) pollen major allergen, Jun a 1."  
 RL J. Allergy Clin. Immunol. 104:608-612(1999).  
 CC -1- DISEASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS  
 CC IN NORTH AMERICA.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 DR EMBL, AF106663; AAD03609.1; -  
 DR EMBL, AF106662; AAD03608.1; -  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec lyase; 1.  
 DR PRINTS; PR00807; AMBALLERGEN.  
 DR Allergen; Glycoprotein; Signal.  
 KW SIGNAL  
 FT CHAIN 1 21  
 FT CARBOHYD 148 148 MAJOR POLLEN ALLERGEN JUN A 1.  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 367 AA; 39824 MW; FC9B81E67565E49 CRC64;

Query Match 21.1%; Score 150; DB 1; Length 367;  
 Best Local Similarity 55.8%; Pred. No. 2.4e-08;  
 Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

-----  
 QY 73 RPLMIIFSGMMNLIKMPMTIAGYKTFDGRRAEVSYYVHNGAKFIRVDGII 124  
 DB 87 KALMIIFSGMMNLIKMPMTIAGYKTFDGRRAEVSYYVHNGAKFIRVDGII 138

RESULT 3  
 MPAL2 CRYUA STANDARD; PRT; 514 AA.  
 AC P43212;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Possible polylacturonase precursor (EC 3.2.1.15) (PG) (Pectinase)  
 DE (Major pollen allergen Cry j 2) (Cry j II).  
 OS Cryptomeria japonica (Japanese cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.  
 OC NCBI\_TaxID=3369;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP TISSUE=Pollen;  
 RX MEDLINE=95010777; PubMed=7926035;  
 RA Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,  
 RA Usui M., Kurimoto M.;  
 RT "Molecular cloning of the second major allergen, Cry j II, from  
 RT Japanese cedar pollen."  
 RL FEBS Lett. 353:124-128(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Pollen;  
 RX MEDLINE=94271186; PubMed=8002972;  
 RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;  
 RT "cDNA cloning and expression of Cry j II the second major allergen of  
 RT Japanese cedar pollen."  
 RL Biochem. Biophys. Res. Commun. 201:1021-1028(1994).  
 RN [3]  
 RP SEQUENCE OF 55-64.  
 RP MEDLINE=90342988; PubMed=2382797;  
 RA Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhasi T.;  
 RT "Identification of the second major allergen of Japanese cedar



RT pollen." ;  
 RL Allergy 45:309-312(1990).  
 CC -1- CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-  
 CC galactosiduronic linkages in pectate and other galacturonans.  
 CC -1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES  
 CC (POLYGALACTURONASES).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D37765; BAA07021.1; -;  
 DR EMBL; D29772; BAA06172.1; -;  
 DR HSPB; P26509; 1BHE.  
 DR InterPro; IPR000743; GH28.  
 DR Pfam; PF00295; Glyco\_hydro\_28; 1.  
 DR PROSITE; PS00502; POLYGALACTURONASE; 1.  
 KM Hydroxylase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;  
 KM Amyloplast; Glycoprotein; Allergen.  
 FT SIGNAL 1 22  
 FT PROPEP 23 45  
 FT CHAIN 46 433  
 FT ACT SITE 434 514  
 FT ACT SITE 278 278  
 FT CARBOHYD 460 460  
 FT CARBOHYD 472 472  
 FT CONFLICT 5 5  
 FT CONFLICT 12 12  
 FT CONFLICT 34 35  
 FT CONFLICT 37 37  
 FT CONFLICT 88 88  
 FT CONFLICT 98 98  
 FT CONFLICT 451 451  
 FT CONFLICT 454 454  
 FT CONFLICT 504 504  
 FT CONFLICT 507 507  
 SQ SEQUENCE 514 AA; 56645 MW; 624611C3FABD502 CRC64;  
 Query Match 20.9%; Score 148.5; DB 1; Length 514;  
 Best Local Similarity 41.4%; Pred. No. 4.9e-08;  
 Matches 41; Conservative 8; Mismatches 21; Indels 29; Gaps 2;  
 QY 31 IDIFASGNFHLQKNTTGTGRWKNRRLWLPALGTGLMGRRLPMTISGNNTKMP 90  
 DB 236 IDIFASGNFHLQKNTTGTG-----DDCVAIGTGSNIVIEDL 272  
 QY 91 MYIAGYKTFDG-----RRAEVSYVHNGAKFIRVDGI 123  
 DB 273 ICGPHGISIGSLGRENSRAEVSYVHNGAKFIDTONGL 311  
 RESULT 4  
 MPAL\_CUPAR STANDARD; PRT; 346 AA.  
 ID 09SCG9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Major pollen allergen Cup a 1.  
 OS Cupressus arizonica.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 RX NCBI\_Taxid=49011;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20571526; PubMed=11122214;  
 RA Aceituno E., Del Pozo V., Minguez A., Arrieta I., Cortegano I.,  
 RA Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.;

RT "Molecular cloning of major allergen from Cupressus arizonica pollen:  
 RT Cup a 1." ;  
 RL Clin. Exp. Allergy 30:1750-1758(2000).  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A II/AMB A II/CRY J I SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ243570; CAB62551.1; -;  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase1.1.  
 DR PRINTS; PRO0807; AMBALLERGEN.  
 KM Allergen; Glycoprotein.  
 FT CARBOHYD 127 127  
 FT CARBOHYD 157 157  
 FT CARBOHYD 272 272  
 SQ SEQUENCE 346 AA; 37589 MW; F1281DCDA1D5DFD0 CRC64;  
 Query Match 20.1%; Score 143; DB 1; Length 346;  
 Best Local Similarity 53.8%; Pred. No. 1.2e-07;  
 Matches 28; Conservative 8; Mismatches 16; Indels 0; Gaps 0;  
 QY 73 RPLMTIFSGNNMILKLMPIYAGYKTFDGRRAEVSYVHNGAKFIRVDGI 124  
 DB 66 KALMTIFSGNNMILKLMPIYAGYKTFDGRGAVALHNGGCEFLMRASHVI 117  
 RESULT 5  
 MPAL\_CHAOB STANDARD; PRT; 375 AA.  
 ID 096385;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Major pollen allergen Cha o 1 precursor.  
 OS Chamaecyparis obtusa (Japanese cypress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;  
 CC Chamaecyparis.  
 RX NCBI\_Taxid=13415;  
 RX [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP TISSUE=Pollen;  
 RX MEDLINE=96265194; PubMed=8676896;  
 RA Suzuki M., Komiyama N., Itoh M., Sone T., Kuno K., Takagi I.,  
 RA Ohta N.;  
 RT "Purification, characterization and molecular cloning of Cha o 1, a  
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen." ;  
 RL Mol. Immunol. 33:451-460(1996).  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D45404; BAA08246.1; -;  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase1.1.  
 DR PRINTS; PRO0807; AMBALLERGEN.  
 KM Allergen; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 22 375  
 MAJOR POLLEN ALLERGEN CHA O 1.

FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF706DDBF CRC64;

Query Match. 20.0%; Score 142; DB 1; Length 375;  
 Best Local Similarity 57.7%; Pred. No. 1.7e-07;  
 Matches 30; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

Qy 73 RLPLWIFSGNNWIKLMPYIAGYKTFDGRRAVSVYHNGAKFRVYDGI 124  
 Db 87 RLPLWIFSGNNWIKLMPYIAGYKTFDGRRAVSVYHNGAKFRVYDGI 138

## RESULT 6

9612\_LYCES STANDARD; PRT; 404 AA.  
 AC P24396;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Style development-specific protein 9612 precursor.  
 GN 9612.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CV\_VF36; TISSUE=Pistill;  
 RX MEDLINE=91117185; PubMed=227637;  
 RA Budelier K.A., Smith A.G., Gasser C.S.;  
 RT "Regulation of a stylar transmitting tissue-specific gene in  
 wild-type and transgenic tomato and tobacco."

CC -1- FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING  
 TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION  
 SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE  
 GROWTH

CC -1- SUBCELLULAR LOCATION: Secreted

CC -1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS  
 FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING

CC TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH  
 LOWER LEVELS IN THE ANTHERS AND VEGETATIVE ORGANS.

CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHERESIS.

CC -1- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P157122)

CC AND P56 (AC P15721).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X55193; GA38979.1; -;  
 DR PIR; S12209; S12209;  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PR00807; AMBALLERGEN.  
 KW Signal.

FT CHAIN 1 20 POTENTIAL.  
 FT CARBOHYD 21 404 STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612.  
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 404 AA; 44298 MW; B26ED69B128D8675 CRC64;

Query Match 14.5%; Score 103; DB 1; Length 404;  
 Best Local Similarity 30.8%; Pred. No. 0.0021;

Matches 33; Conservative 13; Mismatches 31; Indels 30; Gaps 4;

Qy 28 GRRIDFASKNFHLOKMTIGTRKWNRIW-----LQPAKLTGFTLM 70  
 Db 67 GKSPAFSCAIGFGKNAIG-----GKNGRIYVTDGNDPDVNPXKTLRAH-----VI 116

Qy 71 GRRPLWIFSGNNWIKLMPYIAGYKTFDGRRAVSVYHNGAKFI 117  
 Db 117 ODEPLWIFKRDVYQLKQELVMSYKTTIDRGAS---VHISGPCI 160

## RESULT 7

MB1\_AMBAR STANDARD; PRT; 396 AA.  
 AC P27759;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pollen allergen Amb a 1.1 precursor (Antigen B) (Antigen Amb a I).  
 OS Ambrosia artemisiifolia (Short ragweed) Embryophyta; Tracheophyta;  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;  
 OC Heliantheae; Ambrosia.  
 OC NCBI\_TaxID=4212;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

RC TISSUE=Pollen;  
 RX MEDLINE=91093235; PubMed=1702434;  
 RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,  
 RA Klapper D.G.;  
 RT "Cloning of Amb a I (antigen B), the major allergen family of short  
 ragweed pollen."

RT J. Biol. Chem. 266:1229-1236 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANTS.

RC TISSUE=Pollen;  
 RX MEDLINE=92234570; PubMed=1809687;  
 RA Griffith I.J., Pollack J., Klapper D.G., Rogers B.L., Nault A.K.;  
 RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens  
 in Ambrosia artemisiifolia (short ragweed)."

RT Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.

CC -1- PTM: The N-terminus is blocked  
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M63116; - NOT ANNOTATED\_CDS.  
 DR EMBL; M80558; AAA32665.1; -;  
 DR PIR; A39099; A39099.  
 DR PIR; A53240; A53240.  
 DR InterPro; IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PR00807; AMBALLERGEN.  
 DR Antigen; Allergen; Signal; Multigene family; Polymorphism.  
 KW Signal.

FT CHAIN 1 25 POTENTIAL.  
 FT CARBOHYD 26 396 POLLEN ALLERGEN AMB A 1.1.  
 FT VARIANT 92 92 E -> D.  
 SQ SEQUENCE 396 AA; 42709 MW; 0CE7DCEB328841D CRC64;

Query Match 13.9%; Score 98.5; DB 1; Length 396;  
 Best Local Similarity 34.7%; Pred. No. 0.006; Mismatches 22; Indels 15; Gaps 3;

QY 73 RPLMIIFSGNNIKLKMVYIAGYKTFDGRRAEVSYYH-----VNGAK--FIRVD----- 121  
 DB 113 RPLMIIFSRDVIIRLDKEMVYVNSDKTIDGRGAKVYIINAGFTLNGVKNVITHINIMEDVK 172  
 QY 122 -----GIIAAYQNPAS 132  
 DB 173 VNGGLIRKNDGPAA 187  
 RESULT 8  
 MP13\_AMBAR STANDARD; PRT; 397 AA.  
 ID MP13\_AMBAR  
 AC P27761;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pollen allergen Amb a 1.3 precursor (Antigen E) (Antigen Amb a I).  
 OS Ambrosia artemisiifolia (Short ragweed).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;  
 OC Heliantheae; Ambrosia.  
 NCBI\_TaxID=4212;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pollen;  
 RX MEDLINE=91093235; PubMed=1702434;  
 RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,  
 RA Klapper D.G.;  
 RT "Cloning of Amb a I (antigen E), the major allergen family of short  
 RT ragweed pollen.";  
 RL J. Biol. Chem. 266:1229-1236(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANTS.  
 RC TISSUE=Pollen;  
 RX MEDLINE=92234570; PubMed=1809687;  
 RA Griffith I.J., Pollock U., Klapper D.G., Rogers B.L., Nault A.K.;  
 RT "Sequence polymorphism of Amb a I and Amb a II, the major allergens  
 RT in Ambrosia artemisiifolia (short ragweed).";  
 RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: M62961; AAA32668.1; -  
 CC EMBL: M80560; AAA32669.1; ALT\_SEQ.  
 CC PIR: C39099; C39099.  
 CC InterPro: IPR002022; Amb allergen.  
 CC Pfam: PF00544; pec\_lyase; 1.  
 CC PRINTS: PR00807; AMBALLERGEN.  
 CC Antigen, Allergen, Signal, Multigene family, Polymorphism.  
 FT SIGNAL 1 25  
 FT CHAIN 26 397  
 FT VARIANT 48 48 L -> Y.  
 FT SEQUENCE 397 AA; 42928 MW; C8DB41257590DD0A CRC64;  
 Query Match 13.6%; Score 97; DB 1; Length 397;  
 Best Local Similarity 29.5%; Pred. No. 0.0086;  
 Matches 31; Conservative 7; Mismatches 33; Indels 34; Gaps 3;  
 QY 52 WKNNRIWQ-----FAKLT-----GTLMKRRRLMIIF 79

DB 61 WENNRQALADCAQGFAGKGYGKMGVDVYTVTSNLDVDVANPEKGTLLRPAQAQNRLMIIF 120  
 QY 80 SGNNMIKILKMVYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124  
 DB 121 KNDMVTINLNOELVNSDKTIDGRGVKVEI--INGGLTLMNKNII 163  
 RESULT 9  
 PEL\_TOBAC STANDARD; PRT; 397 AA.  
 ID PEL\_TOBAC  
 AC P40572;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Pectate lyase precursor (EC 4.2.2.2).  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 NCBI\_TaxID=4097;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Samsun; TISSUE=Pollen;  
 RX MEDLINE=93043039; PubMed=1421152;  
 RA Rogers H.J., Harvey A., Lonsdale D.M.;  
 RT "Isolation and characterization of a tobacco gene with homology to  
 RT pectate lyase which is specifically expressed during  
 RT microsporogenesis.";  
 RL Plant Mol. Biol. 20:493-502(1992).  
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give  
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at  
 CC their non-reducing ends.  
 CC -1- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN  
 CC DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: X67158; CAA47630.1; -  
 CC EMBL: X67159; CAA47631.1; -  
 CC EMBL: X61102; CAA43414.1; -  
 CC PIR: S26211; S26211.  
 CC InterPro: IPR002022; Amb allergen.  
 CC Pfam: PF00544; pec\_lyase; 1.  
 CC PRINTS: PR00807; AMBALLERGEN.  
 CC Lyase, Signal.  
 FT SIGNAL 1 25  
 FT CHAIN 26 397  
 FT ACT SITE 272 272  
 FT CARBOHYD 134 134  
 FT CARBOHYD 227 227  
 FT CONFLICT 156 156  
 FT CONFLICT 189 189  
 FT CONFLICT 200 200  
 FT CONFLICT 202 202  
 FT CONFLICT 249 249  
 FT SEQUENCE 397 AA; 44351 MW; EFOA82CE5DA7643F CRC64;  
 Query Match 13.1%; Score 93; DB 1; Length 397;  
 Best Local Similarity 42.9%; Pred. No. 0.022;  
 Matches 21; Conservative 7; Mismatches 17; Indels 4; Gaps 2;  
 QY 67 FTLMGRPLMIITSGNNMIKILKMVYIAGYKTFDGRRAEVSYYH-NGA 114  
 DB 105 YGVIOKEPLMIITFGKMKIKLSRELIVTSNKTIDGSGFN---VHIONGA 150

## RESULT 10

MP12 AMBAR STANDARD; PRT; 398 AA.  
 AC P27760;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pollen allergen Amb a 1.2 precursor (Antigen E) (Antigen Amb a I).  
 OS Ambrosia artemisiifolia (Short ragweed).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;  
 OC Heliantheae; Ambrosia.  
 CC NCBI\_TaxID=4212;  
 RX MEDLINE=92234570; PubMed=1809687;  
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;  
 RA Ratnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,  
 RA Klapper D.G.;  
 RT "Cloning of Amb a I (antigen E), the major allergen family of short  
 ragweed pollen.";  
 RT J. Biol. Chem. 266:1229-1236 (1991).  
 CC [2]  
 CC SEQUENCE FROM N.A., AND VARIANTS.  
 CC TISSUE=Pollen;  
 CC MEDLINE=92234570; PubMed=1809687;  
 CC Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;  
 CC "Sequence polymorphism of Amb a I and Amb a II, the major allergens  
 CC in Ambrosia artemisiifolia (short ragweed).";  
 CC Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M62981; AAA32667.1; -;  
 CC EMBL: M60559; AAA32667.1; -;  
 CC PIR: B39099; B39099.  
 CC InterPro: IPR002022; Amb allergen.  
 CC Pfam: PF00544; Pec lyase; 1.  
 CC PRINTS: PR00807; AMBALLERGEN.  
 CC KMW Antigen; Allergen; Signal; Multigene family; Polymorphism.  
 CC SIGNAL 1 25  
 CC CHAIN 26 398 POTENTIAL.  
 CC VARIANT 345 345 POLLEN ALLERGEN AMB A 1.2.  
 CC VARIANT 381 381 R -> K.  
 CC VARIANT 381 381 L -> I.  
 CC SEQUENCE 398 AA; 43664 MW; 020DC662D9B7416C CRC64;

Query Match 13.1%; Score 93; DB 1; Length 398;  
 Best Local Similarity 30.5%; Pred. No. 0.022;  
 Matches 32; Conservative 5; Mismatches 34; Indels 34; Gaps 3;

QY 52 WKNNRIWQ-----FAKLT-----GFTMGRRPILITF 79  
 DB 62 WANNRQALADAOAGFACTGYGKGGDYVYTSKDDVAVPKSGTALFAAQRPLIITF 121  
 QY 80 GGNNNIKIKMPMYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124  
 DB 122 KRNWYIHLNQLVNVSDKITDGRGKVKVI--VNAGLTLMNVKNII 164

## RESULT 11

MPA2 AMBAR STANDARD; PRT; 397 AA.  
 AC P27762;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pollen allergen Amb a 2 precursor (Antigen K) (Antigen Amb a II).  
 OS Ambrosia artemisiifolia (Short ragweed).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;  
 OC Heliantheae; Ambrosia.  
 CC NCBI\_TaxID=4212;  
 RX MEDLINE=92013060; PubMed=1717566;  
 RA Rogers B.L., Morgenstern J.P., Griffith I.J., Yu X.-B.,  
 RA Counsell C.M., Brauer A.W., King T.P., Garmen R.D., Kuo M.-C.C.;  
 RT "Complete sequence of the allergen Amb alpha II. Recombinant  
 RT expression and reactivity with T cells from ragweed allergic  
 RT patients.";  
 RT J. Immunol. 147:2547-2552 (1991).  
 CC [2]  
 CC SEQUENCE FROM N.A., AND VARIANTS.  
 CC TISSUE=Flower;  
 CC MEDLINE=92234570; PubMed=1809687;  
 CC Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;  
 CC "Sequence polymorphism of Amb a I and Amb a II, the major allergens  
 CC in Ambrosia artemisiifolia (short ragweed).";  
 CC Int. Arch. Allergy Appl. Immunol. 96:296-304 (1991).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M80561; AAA32671.1; -;  
 CC PIR: A46469; A46469.  
 CC InterPro: IPR002022; Amb allergen.  
 CC Pfam: PF00544; Pec lyase; 1.  
 CC PRINTS: PR00807; AMBALLERGEN.  
 CC KMW Antigen; Allergen; Signal; Polymorphism.  
 CC SIGNAL 1 25  
 CC CHAIN 26 397 POTENTIAL.  
 CC VARIANT 70 70 N -> D (DETECTED ONLY IN FLOWER DNA).  
 CC VARIANT 138 138 K -> T (DETECTED ONLY IN FLOWER DNA).  
 CC VARIANT 321 321 K -> R (DETECTED ONLY IN FLOWER DNA).  
 CC SEQUENCE 397 AA; 44082 MW; C78617E4C9A3D1DD CRC64;

Query Match 12.9%; Score 92; DB 1; Length 397;  
 Best Local Similarity 42.3%; Pred. No. 0.028;  
 Matches 22; Conservative 7; Mismatches 21; Indels 2; Gaps 1;

QY 73 RPLWIFSGNNNIKIKMPMYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 124  
 DB 114 RPLWIFSGNNNIKIKMPMYIAGYKTFDGRRAEVSYYHNGAKFIRVDGII 163  
 RESULT 12  
 ID P566\_LYCSES  
 AC P15721;  
 STANDARD; PRT; 398 AA.

DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Probable pectate lyase P56 precursor (EC 4.2.2.2).  
 GN LATS6.  
 OS Lycopersicon esculentum (tomato).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv VP36; TISSUE=Anther;  
 RX MEDLINE=91322485; PubMed=1983191;  
 RA Wing R.A., Yamaguchi J., Larebell S.K., Ursin V.M., McCormick S.;  
 RT "Molecular and genetic characterization of two pollen-expressed genes  
 that have sequence similarity to pectate lyases of the plant pathogen  
 Erwinia.";  
 RL Plant Mol. Biol. 14:17-28(1990).  
 RL [2]  
 RP REVISIONS.  
 RC STRAIN=cv VP36; TISSUE=Anther;  
 RA Wing R.A.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE  
 CC GROWTH.  
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give  
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at  
 CC their non-reducing ends.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X15500; CAA3524.1; -  
 CC PIR; S08383; S08383.  
 DR InterPro: IPR002022; Amb\_allergen.  
 DR Pfam: PF00544; pec\_lyase.1.  
 DR PRINTS: PR00807; AMBALLERGEN.  
 KW Lyase; Multigene family; Signal.  
 FT SIGNAL 1 27  
 FT CHAIN 28 398 OR 22 (POTENTIAL).  
 FT ACT SITE 273 273 PROBABLE PECTATE LYASE P56.  
 FT CARBOHYD 135 135 POTENTIAL.  
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 398 AA; 44563 MW; 8D676250BD8C7C8 CRC64;  
 Query Match 12.8%; Score 91; DB 1; Length 398;  
 Best Local Similarity 40.8%; Pred. No. 0.036;  
 Matches 20; Conservative 10; Mismatches 15; Indels 4; Gaps 2;  
 QY 67 FLTMGRPMTIFSGNMNITKMKPMYIAGYKTPDGRRAVSYYHNGA-FIRVDGII 114  
 Db 106 FGVIOGPIMTIFARSMTIRLTRELIVSNKTIIDRG--KXVHLANGA 151  
 RESULT 13  
 MP14\_AMBAR STANDARD; PRT; 392 AA.  
 ID \_MP14\_AMBAR  
 AC P28744;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pollen allergen Amb a 1.4 precursor (Antigen E) (Antigen Amb a 1).  
 OS Ambrosia artemisiifolia (Short ragweed).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

CC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;  
 CC Heliantheae; Ambrosia.  
 OX NCBI\_TaxID=4212;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS.  
 RC TISSUE=Pollen;  
 RX MEDLINE=92234570; PubMed=1809687;  
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;  
 RT "Sequence polymorphism of Amb a 1 and Amb a 1L, the major allergens  
 in Ambrosia artemisiifolia (short ragweed).";  
 RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A 1/AMB A 1I/CRY J I SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M80562; AAA32670.1; -  
 CC PIR; D53240; D53240.  
 DR InterPro: IPR002022; Amb\_allergen.  
 DR Pfam: PF00544; pec\_lyase.1.  
 DR PRINTS: PR00807; AMBALLERGEN.  
 KW Antigen; Allergen; Signal; Multigene family; Polymorphism.  
 FT SIGNAL 1 25  
 FT CHAIN 26 392  
 FT VALINANT 182 188 SHDGPV -> CNDGPV.  
 SQ SEQUENCE 392 AA; 42842 MW; 7B6219C12F365DA9 CRC64;  
 Query Match 12.5%; Score 89; DB 1; Length 392;  
 Best Local Similarity 44.2%; Pred. No. 0.058;  
 Matches 23; Conservative 7; Mismatches 20; Indels 2; Gaps 1;  
 QY 73 RPLMTIFSGNMNITKMKPMYIAGYKTPDGRRAVSYYHNGA-FIRVDGII 124  
 Db 114 RPLMTIFARDWVRRLDRLAINDKTIIDGAVETINAGFA--IYVKNIT 163  
 RESULT 14  
 PEL\_LILLO STANDARD; PRT; 434 AA.  
 ID \_PEL\_LILLO  
 AC P40973;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Pectate lyase precursor (EC 4.2.2.2).  
 OS Lilium longiflorum (Trumpet lily).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; liliopsida; Liliales; Liliaceae;  
 CC Lilium.  
 OX NCBI\_TaxID=4690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv Nellie white; TISSUE=Pollen;  
 RA Kim S.R., Finkel D.J., An G.;  
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give  
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronosyl groups at  
 CC their non-reducing ends.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; 217328; CAA78976.1; -  
 DR EMBL; L18911; AAA33398.1; -  
 DR PIR; S29612; S29612.  
 DR InterPro: IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase.1  
 DR PRINTS; PRO0807; AMBALERGEN.  
 DR Lyase; Signal.  
 KW SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 434 PECTATE LYASE.  
 FT ACT SITE 312 312 POTENTIAL.  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;

Query Match 12.5%; Score 89; DB 1; Length 434;  
 Best Local Similarity 55.9%; Pred. No. 0.064;  
 Matches 19; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

Qy 73 RPLMIFSGNNIKLKMPTIAGKTFDGRRAV 106  
 Db 151 RPLMIFGKSMVIRLQELIINNKTIDGRANV 184

RESULT 15  
 PES9\_LYCES  
 ID PES9\_LYCES STANDARD; PRT; 449 AA.  
 AC P15722;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Probable pectate lyase P59 precursor (EC 4.2.2.2).  
 GN LATS9.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. VF36; TISSUE=anther;  
 RX MEDLINE=91322485; PubMed=1983191;  
 RA Wang R.A., Yamaguchi J., Larabell S.K., Ursin V.M., McCormick S.;  
 RT "Molecular and genetic characterization of two pollen-expressed genes  
 RT that have sequence similarity to pectate lyases of the plant pathogen  
 RT Erwinia.";  
 RL Plant Mol. Biol. 14:17-28(1990).  
 CC -1- FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE  
 CC GROWTH.  
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give  
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enuronyl groups at  
 CC their non-reducing ends.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation on  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X15499; CAA33523.1; -  
 DR PIR; S27098; S27098.  
 DR InterPro: IPR002022; Amb\_allergen.  
 DR Pfam; PF00544; pec\_lyase.1.  
 DR PRINTS; PRO0807; AMBALERGEN.  
 KW Lyase; Multigene family; Signal.  
 FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 449 PROBABLE PECTATE LYASE P59.  
 FT ACT SITE 325 325 POTENTIAL.  
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 449 AA; 50893 MW; 17E3AA13F173B03C CRC64;

Query Match 11.6%; Score 82.5; DB 1; Length 449;  
 Best Local Similarity 22.3%; Pred. No. 0.32;  
 Matches 39; Conservative 13; Mismatches 48; Indels 75; Gaps 7;

Qy 3 VYVAPQFPPNRRVFIKVSNTIIGRRIDIFASKNFHLQKNTIGTGR----- 51  
 Db 57 VTLAENQ--KLDTWKELKKV-----KGTNNSTRGLGKXKXTCGCMVTP 101  
 Qy 52 ---WNNRIWLOFAKLTGFTLMG----- 71  
 Db 102 IDKWRCDENWADNKKLADCAMGFSKAIIGKDGFEYVVDNSDYNPDKGTLRHAVI 161

Qy 72 -RRPLMIFSGNNIKLKMPTIAGKTFDGRRAVYVH-NGA---KFIERY 120  
 Db 162 QKEPLMIFKRGKNTIRLHQEMTMSDKTIDARGV---VHTKGGGITLQYIKNV 213

Search completed: April 23, 2003, 18:05:13  
 Job time : 16 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 23, 2003, 18:04:50 ; Search time 85 Seconds

(without alignments)  
324.827 Million cell updates/sec

Title: US-09-142-524D-3

Perfect score: 711

Sequence: 1 MKYTVAFNPGFNNRVPIKR.....KFIKRVGIIAAYQNPASWK 134

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP xvirus:\*  
16: SP bacteriophage:\*  
17: SP archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188	26.4	374	10	Q8RUR1
2	150	21.1	367	10	Q9LRT2
3	150	21.1	367	10	Q9LRT1
4	150	21.1	367	10	Q9M4S6
5	150	21.1	367	10	Q9M4S4
6	150	21.1	367	10	Q9M4S3
7	150	21.1	367	10	Q9M4S2
8	150	21.1	367	10	Q9M4S1
9	146	20.5	367	10	Q9M4S5
10	144	20.3	367	10	Q9M4S6
11	110.5	15.5	501	10	Q93Z04
12	110.5	15.5	542	10	Q9SV40
13	107.5	15.1	507	10	Q9FY19
14	104.5	14.7	463	10	Q9SCP2
15	100.5	14.1	453	10	Q94LR5
16	99	13.9	472	10	Q9LFP5

17	97	13.6	455	10	Q64510	Q64510 arabidopsis
18	95.5	13.4	181	10	Q23666	Q23666 arabidopsis
19	95.5	13.4	410	10	Q9FMK5	Q9FMK5 arabidopsis
20	95.5	13.4	432	10	Q93Z25	Q93Z25 arabidopsis
21	95.5	13.4	459	10	Q23665	Q23665 arabidopsis
22	95.5	13.4	459	10	Q9M952	Q9M952 arabidopsis
23	95	13.4	459	10	Q9FM66	Q9FM66 arabidopsis
24	95	13.4	452	10	Q9LRM5	Q9LRM5 arabidopsis
25	94.5	13.3	226	10	Q23667	Q23667 arabidopsis
26	94.5	13.3	354	10	Q94001	Q94001 arabidopsis
27	94.5	13.3	398	10	Q9M505	Q9M505 arabidopsis
28	94.5	13.3	404	10	Q9SB71	Q9SB71 arabidopsis
29	94.5	13.3	408	10	Q9CSM8	Q9CSM8 arabidopsis
30	94.5	13.3	431	10	Q23017	Q23017 arabidopsis
31	93.5	13.2	407	10	Q9FXD8	Q9FXD8 arabidopsis
32	93.5	13.2	407	10	Q9SDM4	Q9SDM4 arabidopsis
33	93.5	13.2	408	10	Q8W116	Q8W116 arabidopsis
34	93	13.1	450	10	Q40319	Q40319 medicago sa
35	91.5	12.9	368	10	Q9C8G4	Q9C8G4 arabidopsis
36	90.5	12.7	409	10	Q93XJ1	Q93XJ1 salix gligt
37	90	12.7	409	10	Q9LRT0	Q9LRT0 arabidopsis
38	89.5	12.6	324	10	Q8S345	Q8S345 capsicum an
39	89.5	12.6	398	10	Q43783	Q43783 musa acumin
40	88.5	12.4	297	10	Q9FY87	Q9FY87 arabidopsis
41	88	12.4	438	10	Q43862	Q43862 zea mays (m
42	87	12.2	368	10	Q94FT5	Q94FT5 fragaria an
43	85.5	12.0	416	10	Q9M828	Q9M828 arabidopsis
44	85	12.0	227	10	Q23668	Q23668 arabidopsis
45	84.5	11.9	401	10	Q24554	Q24554 zinnia eleg

## ALIGNMENTS

### RESULT 1

Q8RUR1 PRELIMINARY; PRT; 374 AA.

AC Q8RUR1

DT 01-JUN-2002 (TREMUR1. 21, Created)

DT 01-JUN-2002 (TREMUR1. 21, Last sequence update)

DE Cry j 1 precursor.

GN Cry j 1.1 OR Cry j 1.2.

OS Cryptomeria japonica (Japanese cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.

OK NCBI\_TaxID=3369;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=POLLEN;

RA Futamura N., Shinohara K.;

RT "Isolation and characterization of cDNAs encoding major allergen Cry j

RL 1 from Cryptomeria japonica pollen."

DR EMBL; AB081309; BAB86286.1; -

DR EMBL; AB081310; BAB86287.1; -

KW SIGNAL.

FT SIGNAL. 1 21

FT CHAIN. 22 374

SO SEQUENCE 374 AA; 40720 MW; 90D085D24BF2BD4 CRC64;

Query Match 26.4%; Score 188; DB 10; Length 374;

Best Local Similarity 71.2%; Pred. No. 2.6e-12; Mismatches 11; Indels 0; Gaps 0;

Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

Qy 73 RPLMTISGNMNTIKLKMPIYAGYKTFDGRRAVSIVYVNGAKFIRVDGI 124

Db 87 RPLMTISGNMNTIKLKMPIYAGYKTFDGRRAVSIVYVNGAKFIRVDGI 134

RESULT 2

Q9LRT2 PRELIMINARY; PRT; 367 AA.

```

AC 09L1T2:
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-2.
OS Cupressus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
NCBI_TaxID=33584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21315424; PubMed=11422137;
RA Midoro-Horitsu T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL; AF151427; AAF80164.1;-.
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39768 MW; 0A6AC2F1BAF89586 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNNIKLKMPYIAGYKTFDGRRAEVSYVHVGAKFIRVDGII 124
DB 87 KALMIIFSGNNIKLKMPLYVAGHKITIDRGADVHLGNGPCLFMRKVSHTV 138

RESULT 3
O9L1T1 PRELIMINARY; PRT; 367 AA.
AC 09L1T1:
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Pollen major allergen 1-1.
OS Juniperus virginiana (Eastern red cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
NCBI_TaxID=33584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21315424; PubMed=11422137;
RA Midoro-Horitsu T.M., Goldblum R.M., Brooks E.G.;
RT "Identification of mutations in the genes for the pollen allergens of
RT eastern red cedar (Juniperus virginiana).";
RL Clin. Exp. Allergy 31:771-778(2001).
DR EMBL; AF151429; AAF80166.1;-.
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39708 MW; DCBD1981A74E4711 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNNIKLKMPYIAGYKTFDGRRAEVSYVHVGAKFIRVDGII 124
DB 87 KTLMIIFSGNNIKLKMPLYVAGHKITIDRGADVHLGNGPCLFMRKVSHTV 138

RESULT 4
O9M4S6 PRELIMINARY; PRT; 367 AA.
AC 09M4S6:
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.

```

```

GN CUP81.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RX Monnaie R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257491; AAF72625.1;-.
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39835 MW; B4E9C60108C2C5A3 CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNNIKLKMPYIAGYKTFDGRRAEVSYVHVGAKFIRVDGII 124
DB 87 KALMIIFSGNNIKLKMPLYVAGHKITIDRGADVHLGNGPCLFMRKVSHTV 138

RESULT 5
O9M4S4 PRELIMINARY; PRT; 367 AA.
AC 09M4S4:
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP81.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;
RN [1]
RP SEQUENCE FROM N.A.
RX Monnaie R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RT of Cupressus sempervirens.";
RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257493; AAF72627.1;-.
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39934 MW; 974D3011D74E3D6E CRC64;

Query Match 21.1%; Score 150; DB 10; Length 367;
Best Local Similarity 55.8%; Pred. No. 3.7e-08;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 73 RPLMIIFSGNNIKLKMPYIAGYKTFDGRRAEVSYVHVGAKFIRVDGII 124
DB 87 KALMIIFSGNNIKLKMPLYVAGHKITIDRGADVHLGNGPCLFMRKVSHTV 138

RESULT 6
O9M4S3 PRELIMINARY; PRT; 367 AA.
AC 09M4S3:
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cup s 1 pollen allergen.
GN CUP81.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
NCBI_TaxID=13469;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257494; AAF72628.1;
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39832 MW; B5DFB5A61C07A53 CRC64;

Query Match
Best Local Similarity 21.1%; Score 150; DB 10; Length 367;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMIFSGNNMNIKLKMPYIAGYKTFDGRRAEVSYYHVNAGAKFIRVDGII 124
Db 87 KALMIFSQNNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSIVI 138

RESULT 7
ID Q9M4S2 PRELIMINARY; PRT; 367 AA.
AC Q9M4S2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Cup s 1 pollen allergen.
GN CUPs1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257495; AAF72629.1;
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39819 MW; AE7B055A61C07A53 CRC64;

Query Match
Best Local Similarity 21.1%; Score 150; DB 10; Length 367;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMIFSGNNMNIKLKMPYIAGYKTFDGRRAEVSYYHVNAGAKFIRVDGII 124
Db 87 KALMIFSQNNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSIVI 138

RESULT 8
ID Q93X51 PRELIMINARY; PRT; 367 AA.
AC Q93X51;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, last annotation update)
DE Putative allergen jun o 1.
GN JUN o 1.
OS Juniperus oxycedrus (Prickly juniper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.
OX NCBI_TaxId=69008;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=POLLEN;
RC TISSUE=POLLEN;
RA TISSUE=POLLEN;
RA TISSUE=POLLEN;
RT "Cloning of Juniperus oxycedrus major allergen.";

```

```

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ293767; CAC48400.1;
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
SQ SEQUENCE 367 AA; 39808 MW; 5D28204DBDFD1B9D7 CRC64;

Query Match
Best Local Similarity 21.1%; Score 150; DB 10; Length 367;
Matches 29; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMIFSGNNMNIKLKMPYIAGYKTFDGRRAEVSYYHVNAGAKFIRVDGII 124
Db 87 KALMIFSQNNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSIVI 138

RESULT 9
ID Q9M4S5 PRELIMINARY; PRT; 367 AA.
AC Q9M4S5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Cup s 1 pollen allergen.
GN CUPs1.
OS Cupressus sempervirens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=13469;
RN [1]
RP SEQUENCE FROM N.A.
RA Monsalve R.I., Villalba M., Rodriguez R.;
RT "Cloning and expression of Cup s 1, the major allergen of the pollen
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257492; AAF72626.1;
DR InterPro; IPR020222; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
SQ SEQUENCE 367 AA; 39894 MW; 5D56FC0E3263B741 CRC64;

Query Match
Best Local Similarity 20.5%; Score 146; DB 10; Length 367;
Matches 28; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 73 RPLMIFSGNNMNIKLKMPYIAGYKTFDGRRAEVSYYHVNAGAKFIRVDGII 124
Db 87 KALMIFSQNNMNIKLKMPLYVAGHKTIDGRGADVHLGNGGPCLFMRKVSIVI 138

RESULT 10
ID Q93XL6 PRELIMINARY; PRT; 367 AA.
AC Q93XL6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-JUN-2002 (TREMBlrel. 21, last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, last annotation update)
DE Putative allergen Cup a 1 precursor.
GN CUP A 1.
OS Cupressus arizonica.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.
OX NCBI_TaxId=49011;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=POLLEN;
RC TISSUE=POLLEN;
RA Buterom C., Di Felice G., Pini C.;
RT "Cloning of Cupressus arizonica major allergen.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278498; CAC37790.2;
KW Signal.
FT SIGNAL 1 21
FT CHAIN 22 367
SQ SEQUENCE 367 AA; 39809 MW; AFE97260423A9F28 CRC64;

```

Db 237 IDIFASKRFEIEKCTIGT-----DDCVAVGIGSSNLTIKWL

Db 237 IDIFASKRFEIEKCTIGT-----DDCVAVGIGSSNLTIKWL

QY 91 MYAGYKTFD-----RAEVSYHVNGAKFIRVDGI 123  
 DB 274 TCGPBGMSIGLKGKNSRSEVSFVHLDGAKFIDTQNGL 312

## RESULT 14

09SCP2 PRELIMINARY; PRT; 463 AA.  
 ID 09SCP2;  
 AC 09SCP2;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE Pectate lyase-like protein.  
 GN T4D2.120.

OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Nykatura G., Fartmann B., Dauner D., Sterr W., Holland R.,  
 RA Weichselgartner M., Mewes H.W., Lemcke K., Mayer K.F.X., Quettier F.,  
 RA Salanoubat M.;  
 RL Submitted (NCV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL132958; CAB64222.1;  
 DR InterPro: IPR02022; Amb\_allergen.  
 DR Pfam: PF00544; pec\_lyase; 1.  
 DR PRINTS: PR00807; AMBALERGEN.  
 KW Lyase.

SO SEQUENCE 463 AA; 51225 MW; 3424B0DE4ADBA604 CRC64;

Query Match 14.7%; Score 104.5; DB 10; Length 463;  
 Best Local Similarity 30.5%; Pred. No. 0.0047;  
 Matches 32; Conservative 8; Mismatches 32; Indels 33; Gaps 3;

QY 52 WKNRRIML-----QPAKLTLG-----FTLMGRRLMITIP 79  
 DB 66 WPNNRQGLADCGIGFGYALGKGQGYFVTDSSDDANVPKQTLRYGVIGSEPLMTIV 125

QY 80 SGNNNITLKMPTIAGYKTFDGRRAEVSYHVNGAKFIRVDGII 124  
 DB 126 PSNNMIRLKOELLFNSYKTLDRGANVHIVG-GGCITLQYVSNII 169

## RESULT 15

094LRS PRELIMINARY; PRT; 453 AA.  
 ID 094LRS;  
 AC 094LRS;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Putative pectate lyase.  
 OS Oryza sativa (Rice)  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Eriocaridaceae; Oryzaeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]

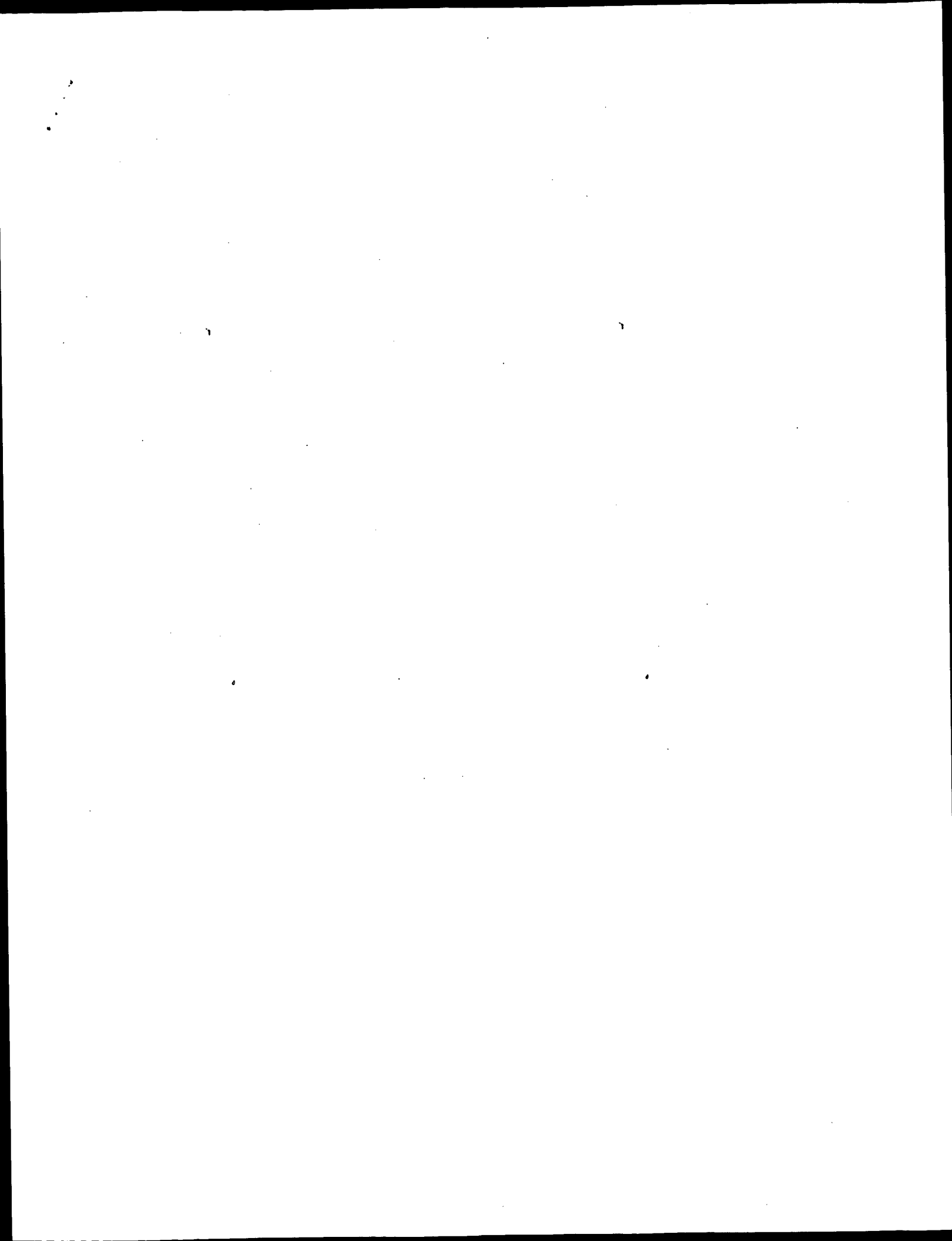
RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
 RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tselirin T.,  
 RA Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., Vanaken S.B.,  
 RA Uteback T.R., Feldblum T.V., Quackenbush J., Salzberg S.L.,  
 RA White O., Frazer C.M.;  
 RT "Oryza sativa chromosome 10 BAC OSUNB0011a08 genomic sequence.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC034258; AAK54283.1; ..

DR InterPro: IPR002022; Amb\_allergen.  
 DR Pfam: PF00544; pec\_lyase; 1.  
 KW Lyase.  
 SO SEQUENCE 453 AA; 48006 MW; 1411BBE1A40901DC CRC64;

Query Match 14.1%; Score 100.5; DB 10; Length 453;  
 Best Local Similarity 48.8%; Pred. No. 0.013;  
 Matches 20; Conservative 7; Mismatches 13; Indels 1; Gaps 1;

QY 74 PLMTIFSGNNNITLKMPTIAGYKTFDGRRAEVSYHVNGA 114  
 DB 140 PLMTIVFAGDNTIRLNEBLVNSYKTLDRGANV-HVGAGCA 179

Search completed: April 23, 2003, 18:07:08  
 Job time : 89 secs



Japanese cedar pollen  
Cry J I Japanese C  
Cedar allergen 4930  
Japanese cedar all  
Cedar pollen aller  
Cedar pollen aller  
Cedar pollen aller  
Cedar pollen aller  
Jun s I. Juniperu  
Jun v I. Juniperu  
Japonicum allergen  
Japonicum allergen  
Japanese cedar pol  
Japanese cedar pol  
Japonicum allergen  
Japonicum allergen  
Japan cedar pollen  
Japan cedar pollen  
Cedar pollen aller  
Japanese cedar all  
Japanese cedar all  
Cedar pollen aller  
Chamaecyparis obtu  
Japanese cypress p  
Chamaecyparis obtu  
Chamaecyparis aller  
Cedar pollen aller  
T-cell epitope pep  
Artificial sequenc  
Artificial sequenc  
Artificial sequenc  
Artificial sequenc  
Cedar pollen aller  
Artificial sequenc  
Artificial sequenc  
Cedar pollen aller  
Cedar pollen aller  
Chamaecyparis obtu  
Japanese cypress p  
bugi allergen prot

PS Claim 6; Page 32; 58pp; Japanese.  
 CC The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.  
 XX Sequence 134 AA;  
 SQ  
 Query Match 100.0%; Score 711; DB 18; Length 134;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-82;  
 Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MKVTVAFNQGPNNRVRVFKKVSNTIIGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60  
 DB 1 MKVTVAFNQGPNNRVRVFKKVSNTIIGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60  
 QY 61 FAKLTGFTLMGRRLMTIIFSGNNMIKMKMPYIAGYKTFDGRBAEVSYYHNGAKFIRRY 120  
 DB 61 FAKLTGFTLMGRRLMTIIFSGNNMIKMKMPYIAGYKTFDGRBAEVSYYHNGAKFIRRY 120  
 QY 121 DGIITAAVQNPASWK 134  
 DB 121 DGIITAAVQNPASWK 134  
 Db 121 DGIITAAVQNPASWK 134  
 RESULT 2  
 AAM27370  
 ID AAM27370 standard; peptide; 105 AA.  
 AC AAM27370;  
 AC AAM27370;  
 DT 24-MAR-1998 (first entry)  
 DE Multi-epitope peptide used as immunotherapeutic agent #2.  
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 OS Synthetic.  
 OS Synthetic.  
 PN WO9732600-A1.  
 PN WO9732600-A1.  
 PD 12-SEP-1997.  
 PF 10-MAR-1997; 97WO-JP00740.  
 PR 10-MAR-1996; 96JP-0080702.  
 PR 10-MAR-1996; 96JP-0080702.  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;  
 PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;  
 DR WPI; 1997-470495/43.  
 DR WPI; 1997-470495/43.  
 PT Peptide immuno:therapeutic agent to treat allergic diseases -  
 PT contains multi-epitope peptide containing T cell epitope regions  
 PT from different allergens  
 PT from different allergens  
 PS Claim 6; Page 31; 58pp; Japanese.  
 PS Claim 6; Page 31; 58pp; Japanese.  
 XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.  
 CC allergen; have at least 70% reactivity with lymphocytes from patients

CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.  
 XX Sequence 105 AA;  
 SQ  
 Query Match 73.6%; Score 523.5; DB 18; Length 105;  
 Best Local Similarity 78.4%; Pred. No. 9.8e-59;  
 Matches 105; Conservative 0; Mismatches 0; Indels 29; Gaps 2;  
 QY 1 MKVTVAFNQGPNNRVRVFKKVSNTIIGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60  
 DB 1 MKVTVAFNQGPNNRVRVFKKVSNTIIGRRIDIFASKNFHLQKNTTGTGRMKNNRIWQ 60  
 QY 61 FAKLTGFTLMGRRLMTIIFSGNNMIKMKMPYIAGYKTFDGRBAEVSYYHNGAKFIRRY 120  
 DB 61 FAKLTGFTLMGRRLMTIIFSGNNMIKMKMPYIAGYKTFDGRBAEVSYYHNGAKFIRRY 120  
 QY 121 DGIITAAVQNPASWK 134  
 DB 92 DGIITAAVQNPASWK 105  
 Db 92 DGIITAAVQNPASWK 105  
 RESULT 3  
 AAM27369  
 ID AAM27369 standard; peptide; 80 AA.  
 AC AAM27369;  
 AC AAM27369;  
 DT 24-MAR-1998 (first entry)  
 DE Multi-epitope peptide used as immunotherapeutic agent #1.  
 DE Multi-epitope peptide; immunotherapeutic agent; allergic disease;  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 KM T-cell epitope region; allergen; lymphocyte; immunoglobulin E.  
 OS Synthetic.  
 OS Synthetic.  
 PN WO9732600-A1.  
 PN WO9732600-A1.  
 PD 12-SEP-1997.  
 PF 10-MAR-1997; 97WO-JP00740.  
 PR 10-MAR-1996; 96JP-0080702.  
 PR 10-MAR-1996; 96JP-0080702.  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;  
 PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;  
 DR WPI; 1997-470495/43.  
 DR WPI; 1997-470495/43.  
 PT Peptide immuno:therapeutic agent to treat allergic diseases -  
 PT contains multi-epitope peptide containing T cell epitope regions  
 PT from different allergens  
 PT from different allergens  
 PS Claim 6; Page 31; 58pp; Japanese.  
 PS Claim 6; Page 31; 58pp; Japanese.  
 XX The present sequence represents a multi-epitope peptide which is used as  
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2  
 CC or more different allergens (preferably linked via arginine or lysine  
 CC dimers), where the T cell epitope regions have a positivity index  
 CC greater than 100 as measured in a patient group responding to the  
 CC allergen; have at least 70% reactivity with lymphocytes from patients  
 CC responding to the allergen; and are not reactive with immunoglobulin E  
 CC (IgE) antibodies from patients responsive to the allergen. The agent can  
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by  
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.  
 CC allergen; have at least 70% reactivity with lymphocytes from patients

Query Match

45.0%; Score 320; DB 18; Length 80;

```

Qy      92 YIAGKTFDGRRAEVSIVHVGAKFIRRVGDIIAAQNPAWK 134
Db      113 -----DPR-----GITAAIQNPASWK 128

RESULT 5
AAR75388
ID AAR75388 standard; protein; 353 AA.
XX
AC AAR75388;
XX
DT 12-MAR-1996 (first entry)
XX
DE Japanese cedar pollen allergen Cryj I.
XX
KW Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
KM prevention; treatment; cryptomeria pollinosis.
XX
OS Cryptomeria japonica.
XX
FH Key Location/Qualifiers
FT Peptide 61..75 /note="T-cell epitope peptide"
FT FT 91..105 /note="T-cell epitope peptide"
FT Peptide 106..120 /note="T-cell epitope peptide"
FT FT 146..160 /note="T-cell epitope peptide"
FT Peptide 211..225 /note="T-cell epitope peptide"
FT Peptide 326..340 /note="T-cell epitope peptide"
FT Peptide 335..346 /note="T-cell epitope peptide"
FT Peptide
FT Peptide
FT Peptide
PN JP07118295-A.
XX
PD 09-MAY-1995.
XX
PF 20-OCT-1993; 93JP-0262626.
XX
PR 20-OCT-1993; 93JP-0262626.
XX
PA (MEIP ) MEITI MILK PROD CO LTD.
XX
DR WPI; 1995-203834/27.
XX
PT New cryptomeria pollen allergen T-cell epitope peptide - used for
PT prevention, treatment and investigation of Japanese cedar pollinosis
PT
XX
PS Disclosure; Figs 1-2; bpp; Japanese.
CC AAR75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC T-cell epitope peptides AAR89289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
XX
SQ Sequence 353 AA;

Query Match 26.4%; Score 188; DB 16; Length 353;
Best Local Similarity 71.2%; Pred.No.3.2e-15;
Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0.

Oy      73 RPLMIIFSGNMNKKLMPYITIGYCTPFGRAEVSIVHVGAKFIRRVDDGI 124
Db      66 RPLMIIFSGNMNKKLMPYITIGYCTPFGRAEQVYIYGNGPCVFIRKVSNI 117

```

```

XX AC AAR81587;
XX XX
XX DT 24-MAY-1996 (first entry)
XX DE Cedar pollen allergen B.
XX KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
XX KW antibody; pollinosis; therapy; immunotherapy.
XX OS Cryptomeria japonica.
XX PN EP700929-A2.
XX PD 13-MAR-1996.
XX PF 08-SEP-1995; 95SEP-0306295.
XX PR 14-JUL-1995; 95JUL-0200221.
XX PR 10-SEP-1994; 94SEP-0242137.
XX PR 14-JUL-1995; 95JUL-0200204.
XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
XX PI Hino K, Saito S, Taniguchi Y;
XX DR WPI; 1996-140976/15.
XX PT New peptide(s) derived from cedar pollen allergens - activate
XX PT allergen-specific T-cells, but not allergen-specific IgE antibodies,
XX PT used for treating cedar pollinosis
XX PS Claim 5; Page 31-32; 36pp; English.
XX CC Synthetic peptides based on portions of cedar pollen allergens A
XX CC (AAR81586) and B (AAR81587) were tested for their ability to activate
XX CC cedar allergen-specific T-cells, but not allergen-specific IgE
XX CC antibodies. 6 Peptides (AAR81580-R81585) were identified as T-cell
XX CC epitopes. These peptides, plus subsequences (AAR81573-79) essential
XX CC for T-cell recognition, and homologous peptides (AAR81588-96) can
XX CC be used as immunotherapeutic agents to treat or prevent cedar
XX CC pollinosis, avoiding side-effects such as anaphylaxis.
XX SQ Sequence 353 AA;
XX
XX Query Match 26.4%; Score 188; DB 17; Length 353;
XX Best Local Similarity 71.2%; Pred. No. 3.2e-15;
XX Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
OY 73 RPLWTFSGNNMNIKLMPMYTAGYKTFDGRRAEVSYYHVNGAKFRRVDGI 124
DB 66 RPLWTFSGNNMNIKLMPMYTAGYKTFDGRGAQYIYINGGCVFIKRVSNVI 117

RESULT 7
XX ID AAY25664 standard; protein; 373 AA.
XX AC AAY25664;
XX DT 30-SEP-1999 (first entry)
XX DE Cedar allergen 493634 Cry j IB protein fragment.
XX KW Major histocompatibility complex; class II; desensitizing; human;
XX KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
XX KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX OS Cedrus sp.
XX XX

```

```

XX PN W09934826-A1.
XX XX
XX PD 15-JUL-1999.
XX XX
XX DT 11-JAN-1999; 99WO-GB00080.
XX PF 21-SEP-1998; 98GB-0020474.
XX PR 09-JAN-1998; 98GB-0000445.
XX PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.
XX PI Kay AB, Larche M;
XX DR WPI; 1999-458255/38.
XX PT Desensitizing patients to polypeptide allergens
XX PS Example 6; Page 73; 117pp; English.
XX CC This invention describes a novel method of desensitizing a patient to a
XX CC polypeptide allergen and comprises administering to the patient a peptide
XX CC derived from the allergen where restriction to a MHC Class II molecule
XX CC possessed by the patient can be demonstrated for the peptide and the
XX CC peptide is able to induce a late phase response in an individual who
XX CC possesses the MHC Class II molecule. The methods can be used for
XX CC desensitizing patients to allergens present in e.g. grass, tree and weed
XX CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,
XX CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit
XX CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,
XX CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of
XX CC Tenebrio molitor beetle, mammals such as cat, dog, horse, cow, pig,
XX CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to
XX CC produce immunological vaccines which may be used to prevent and/or treat
XX CC conditions involving hypersensitivity to allergens. This sequence
XX CC represents a cedar (Cedrus sp.) allergen 493634 Cry j IB.
XX SQ Sequence 373 AA;
XX
XX Query Match 26.4%; Score 188; DB 20; Length 373;
XX Best Local Similarity 71.2%; Pred. No. 3.5e-15;
XX Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;
OY 73 RPLWTFSGNNMNIKLMPMYTAGYKTFDGRRAEVSYYHVNGAKFRRVDGI 124
DB 86 RPLWTFSGNNMNIKLMPMYTAGYKTFDGRGAQYIYINGGCVFIKRVSNVI 137

RESULT 8
XX ID AAY25668 standard; protein; 373 AA.
XX AC AAY25668;
XX DT 30-SEP-1999 (first entry)
XX DE Japanese cedar allergen 541803 Cry j I precursor protein fragment.
XX KW Major histocompatibility complex; class II; desensitizing; human;
XX KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;
XX KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;
XX KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;
XX KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;
XX KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.
XX OS Cedrus sp.
XX PN W09934826-A1.
XX DT 15-JUL-1999.
XX PF 11-JAN-1999; 99WO-GB00080.
XX PR 21-SEP-1998; 98GB-0020474.
XX XX

```



PR 09-JAN-1998; 98GB-0000445.  
 XX  
 PA (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.  
 XX  
 PI Kay AB, Larche M;  
 XX  
 DR WPI; 1999-458255/38.  
 XX  
 PT Deesensitizing patients to polypeptide allergens  
 XX  
 PS Example 6; Page 75; 117pp; English.  
 XX  
 CC This invention describes a novel method of deesensitizing a patient to a  
 CC polypeptide allergen and comprises administering to the patient a peptide  
 CC derived from the allergen where restriction to a MHC Class II molecule  
 CC possessed by the patient can be demonstrated for the peptide and the  
 CC peptide is able to induce a late phase response in an individual who  
 CC possesses the MHC Class II molecule. The methods can be used for  
 CC desensitizing patients to allergens present in e.g. grass, tree and weed  
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,  
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit  
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,  
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of  
 CC Teribrio molitor beetle, mammals such as cat, dog, horse, cow, pig,  
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to  
 CC produce immunological vaccines which may be used to prevent and/or treat  
 CC conditions involving hypersensitivity to allergens. This sequence  
 CC represents a Japanese cedar (Cedrus sp.) allergen 541803 Cry j I  
 CC precursor.  
 XX  
 SQ Sequence 373 AA;  
 XX  
 Query Match 26.4%; Score 188; DB 20; Length 373;  
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;  
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
 XX  
 QY 73 RPLWIIIFSGNNMIKIKMPMYIAGYKTFDGRRAEVSYYVHNGAKFIRVDGII 124  
 Db 86 RPLWIIIFSGNNMIKIKMPMYIAGYKTFDGRGAQYVYIGNGPGCVFIIKRVSNVI 137  
 XX  
 RESULT 9  
 AAR31937  
 ID AAR31937 standard; Protein; 374 AA.  
 XX  
 AC AAR31937;  
 XX  
 DT 03-JUN-1993 (first entry)  
 XX  
 DE Cry j I.  
 XX  
 KM Japanese cedar pollen; allergen; antigen; allergy; B cell; T cell.  
 XX  
 OS Cryptomeria japonica.  
 XX  
 Key Location/Qualifiers  
 FH Peptide 1..21  
 FT /note= "signal peptide"  
 FT 22..374  
 FT Protein /note= "mature Cry j I"  
 XX  
 MO9301213-A.  
 PN  
 PD 21-JAN-1993.  
 XX  
 PF 10-JUL-1992; 92WO-US05661.  
 XX  
 PR 12-JUL-1991; 91US-0729134.  
 PR 15-JUL-1991; 91US-0730452.  
 XX  
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.  
 XX  
 PI Bond JF, Griffith IU, Pollock J;

XX  
 DR WPI; 1993-045434/05.  
 DR- N-PSDB; AAQ5304.  
 XX  
 PT Nucleic acid sequence encoding Cryptomeria japonica allergen  
 PT for the diagnosis treatment and prevention of allergic reactions  
 PT to Japanese cedar pollen  
 XX  
 PS Claim 11; Page 42; 69pp; English.  
 XX  
 CC Fresh pollen and staminate cone samples were collected from a single  
 CC Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used  
 CC to synthesize cDNA. The cDNA was subjected to successive rounds of  
 CC PCR to yield a full length Cry j I clone. Cry j I or an antigenic  
 CC fragment of it may be used for detecting, treating and preventing an  
 CC allergic response to Japanese cedar pollen allergen. It is capable of  
 CC modifying both the B and T cell response to Cry j I and T cell response  
 CC to a Cry j I antigen.  
 XX  
 SQ Sequence 374 AA;  
 XX  
 Query Match 26.4%; Score 188; DB 14; Length 374;  
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;  
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
 XX  
 QY 73 RPLWIIIFSGNNMIKIKMPMYIAGYKTFDGRRAEVSYYVHNGAKFIRVDGII 124  
 Db 87 RPLWIIIFSGNNMIKIKMPMYIAGYKTFDGRGAQYVYIGNGPGCVFIIKRVSNVI 138  
 XX  
 RESULT 10  
 AAR45541  
 ID AAR45541 standard; Protein; 374 AA.  
 XX  
 AC AAR45541;  
 XX  
 DT 13-JUL-1994 (first entry)  
 XX  
 DE Cry j I pollen allergen.  
 XX  
 KM Japanese cedar; detection; allergy; treatment; diagnosis;  
 KM T cell epitope; sensitivity.  
 XX  
 OS Cryptomeria japonica.  
 XX  
 Key Location/Qualifiers  
 FH Peptide 1..21  
 FT /note= "signal peptide"  
 FT 22..374  
 FT Peptide /note= "mature peptide"  
 XX  
 MO9401560-A.  
 PN  
 PD 20-JAN-1994.  
 XX  
 PF 15-JAN-1993; 93WO-US00139.  
 XX  
 PR 01-SEP-1992; 92US-0938990.  
 PR 10-JUL-1992; 92WO-US05661.  
 XX  
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.  
 XX  
 PI Bond JF, Garman RD, Griffith IU, Kuo M, Pollock J;  
 XX  
 DR WPI; 1994-035066/04.  
 DR N-PSDB; AAQ55271.  
 XX  
 PT Antigens derived from Japanese cedar pollen allergen Cry j I -  
 PT contain at least two T cell epitope(s), used to treat or diagnose  
 PT allergy  
 XX  
 PS Disclosure, Fig 4; 137pp; English.

CC The sequence is that of the Japanese cedar pollen allergen  
 CC Cry j 1 which contains at least two T cell epitopes. Peptide  
 CC antigens derived from it can be used for the treatment and  
 CC diagnosis of allergies associated with Japanese cedar pollen.  
 CC The peptides have enhanced therapeutic properties but reduced  
 CC side effects compared to naturally occurring allergens.

XX  
 XX  
 SQ Sequence 374 AA;

Query Match 26.4%; Score 188; DB 15; Length 374;  
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;  
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 73 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRRAEVSYYHVNCAKIFRRVDGII 124  
 DB 87 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRGAQYVYIGNGPCVFIRKVSNI 138

RESULT 11

AA60166  
 ID AAR60166 standard; Protein; 374 AA.

XX  
 AC AAR60166;

DT 24-MAR-1995 (first entry)

XX Japanese cedar pollen antigen CryjI.

KW Japanese cedar; pollen antigen; allergen; CryjI; sugi; pollinosis.

OS Crytomeria japonica.

XX  
 FH Key location/Qualifiers

FT Peptide 1..21

FT Protein /label= signal\_peptide

FT /label= mature\_CryjI

XX JP06197768-A.

XX 19-JUL-1994.

XX 07-JAN-1993; 93JP-0001116.

XX 07-JAN-1993; 93JP-0001116.

XX (MEIJ) MEIJ SEIKA KAISHA.

XX WPI; 1994-268680/33.

XX N-PSDB; AAQ71601.

XX Sugi (Japanese cedar) pollen antigen CryjI - is useful for

XX diagnosis, treatment and prevention of sugi pollinosis

XX Claim 2; Page 5-7; 9pp; Japanese.

XX The coding sequence for the Japanese cedar ("sugi") pollen allergen

XX CryjI was isolated from a cDNA library prepared from polyA mRNA. All

XX or part of the CryjI protein can be used for diagnosis, treatment

XX and prevention of sugi pollinosis.

XX SQ Sequence 374 AA;

OY 73 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRRAEVSYYHVNCAKIFRRVDGII 124

DB 87 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRGAQYVYIGNGPCVFIRKVSNI 138

RESULT 12

AA82490  
 ID AAR82490 standard; Protein; 374 AA.

XX  
 AC AAR82490;

DT 15-APR-1996 (first entry)

XX Cry j I Japanese Cedar pollen allergen.

KW Cry j I; Japanese cedar pollen allergen; modified; drug production;

XX allergy; Crytomeria japonica.

XX Crytomeria japonica.

XX MO9527786-A1.

XX 19-OCT-1995.

XX 06-APR-1995; 95MO-US04249.

XX 06-DEC-1994; 94US-0350225.

XX 08-APR-1994; 94US-0226248.

XX (TMMU-) IMMUNOLOGIC PHARM CORP.

XX Chen X, Evans S, Franzen HM, Kuo M, Powers SP;

XX Shaked Z;

XX WPI; 1995-366391/47.

XX N-PSDB; AAT04248.

XX Modified Crytomeria japonica (Cry j) I peptide(s) - useful for

XX treating allergy to Japanese cedar pollen allergen or

XX immunologically cross reactive allergens

XX Disclosure; Figure 1; 60pp; English.

XX Novel peptides of cry j I have been modified as a part of a

XX preformulation scheme to develop an optimised drug product for

XX therapeutic treatment of humans suffering from allergy to Japanese

XX cedar pollen allergen or an allergen which is immunologically cross

XX reactive with Japanese cedar pollen allergen. Such modified peptides

XX possess certain characteristics which render them particularly

XX suitable for drug product formulation. Peptide fragments of Cry j I,

XX modified and unmodified, are given in AAR82491-R82525.

XX SQ Sequence 374 AA;

OY 73 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRRAEVSYYHVNCAKIFRRVDGII 124

DB 87 RPLMTIFSGNNMNIKLMPTIAGYKTFDGRGAQYVYIGNGPCVFIRKVSNI 138

RESULT 13

AA25665  
 ID AAY25665 standard; Protein; 374 AA.

XX  
 AC AAY25665;

DT 30-SEP-1999 (first entry)

XX Cedar allergen 493632 Cry j IA protein fragment.

KW Major histocompatibility complex; class II; desensitising; human;

KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;

KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;

KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;

KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;

KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.

OS Cedrus sp.  
 XX MO9934826-A1.  
 PN 15-JUL-1999.  
 PD 11-JAN-1999; 99WO-GB00080.  
 PF 21-SEP-1998; 98GB-0020474.  
 PR 09-JAN-1998; 98GB-0000445.  
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.  
 PA Kay AB, Larche M;  
 PI WPI; 1999-458255/38.  
 DR Desensitizing patients to polypeptide allergens  
 XX Example 6; Page 73; 117pp; English.  
 PS This invention describes a novel method of desensitizing a patient to a  
 CC polypeptide allergen and comprises administering to the patient a peptide  
 CC derived from the allergen where restriction to a MHC Class II molecule  
 CC possessed by the patient can be demonstrated for the peptide and the  
 CC peptide is able to induce a late phase response in an individual who  
 CC possesses the MHC Class II molecule. The methods can be used for  
 CC desensitizing patients to allergens present in e.g. grass, tree and weed  
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,  
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit  
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,  
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of  
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,  
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to  
 CC produce immunological vaccines which may be used to prevent and/or treat  
 CC conditions involving hypersensitivity to allergens. This sequence  
 CC represents a cedar (Cedrus sp.) allergen 493632 Cry j IA.  
 XX  
 SQ Sequence 374 AA;  
 Query Match 26.4%; Score 188; DB 20; Length 374;  
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;  
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
 QY 73 RPLWTFSGNNNITKLKMPYIAGYKTFDGRRAEVSYYHVNGAKFRRVDGI 124  
 DB 87 RPLWTFSGNNNITKLKMPYIAGYKTFDGRGAQYVYIONGPCVPIKRVSNVI 138  
 RESULT 14  
 ID AAY25669 standard; protein; 374 AA.  
 AC AAY25669;  
 XX 30-SEP-1999 (first entry)  
 DT Japanese cedar allergen 541802 Cry j I precursor protein fragment.  
 DE  
 XX Major histocompatibility complex; class II; desensitizing; human;  
 KW allergen; grass; tree; weed; pollen; fungi; mould; food; insect; sting;  
 KW chironomidae; spider; mite; housefly; fruit fly; sheep blow fly; honeybee;  
 KW screw worm fly; grain weevil; silkworm; bee moth; larvae; mealworm; cat;  
 KW cockroach; beetle; dog; horse; cow; pig; sheep; rabbit; rat; guinea pig;  
 KW mice; gerbil; vaccine; treatment; prevention; hypersensitivity; cedar.  
 XX  
 OS Cedrus sp.  
 XX MO9934826-A1.  
 PN 15-JUL-1999.

PF 11-JAN-1999; 99WO-GB00080.  
 XX 21-SEP-1998; 98GB-0020474.  
 PR 09-JAN-1998; 98GB-0000445.  
 XX (IMCO-) IMPERIAL COLLEGE INNOVATIONS LTD.  
 PA Kay AB, Larche M;  
 PI WPI; 1999-458255/38.  
 DR Desensitizing patients to polypeptide allergens  
 XX Example 6; Page 75; 117pp; English.  
 PS This invention describes a novel method of desensitizing a patient to a  
 CC polypeptide allergen and comprises administering to the patient a peptide  
 CC derived from the allergen where restriction to a MHC Class II molecule  
 CC possessed by the patient can be demonstrated for the peptide and the  
 CC peptide is able to induce a late phase response in an individual who  
 CC possesses the MHC Class II molecule. The methods can be used for  
 CC desensitizing patients to allergens present in e.g. grass, tree and weed  
 CC (including ragweed) pollens, fungi and moulds, foods, stinging insects,  
 CC the chironomidae (non-biting midges), spiders and mites, housefly, fruit  
 CC fly, sheep blow fly, screw worm fly, grain weevil, silkworm, honeybee,  
 CC non-biting midge larvae, bee moth larvae, mealworm, cockroach, larvae of  
 CC Tenbrio molitor beetle, mammals such as cat, dog, horse, cow, pig,  
 CC sheep, rabbit, rat, guinea pig, mice or gerbil. They can also be used to  
 CC produce immunological vaccines which may be used to prevent and/or treat  
 CC conditions involving hypersensitivity to allergens. This sequence  
 CC represents a Japanese cedar (Cedrus sp.) allergen 541802 Cry j I  
 XX  
 SQ Sequence 374 AA;  
 Query Match 26.4%; Score 188; DB 20; Length 374;  
 Best Local Similarity 71.2%; Pred. No. 3.5e-15;  
 Matches 37; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
 QY 73 RPLWTFSGNNNITKLKMPYIAGYKTFDGRRAEVSYYHVNGAKFRRVDGI 124  
 DB 87 RPLWTFSGNNNITKLKMPYIAGYKTFDGRGAQYVYIONGPCVPIKRVSNVI 138  
 RESULT 15  
 ID AAB69103 standard; protein; 210 AA.  
 AC AAB69103;  
 XX 23-APR-2001 (first entry)  
 DT Cedar pollen allergen T cell epitope derived protein SBQ ID NO:12.  
 DE  
 XX Japanese cedar; Cryptomeria japonica; cedar pollen allergen;  
 KW T cell epitope; antiasthma; pollinosis.  
 KW  
 XX Cryptomeria japonica.  
 OS Synthetic.  
 XX JP2000327699-A.  
 PN 28-NOV-2000.  
 PD 15-MAR-2000; 2000JP-0071710.  
 PF 15-MAR-1999; 99JP-0068316.  
 PR (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.  
 PA (SANY ) SANKYO CO LTD.  
 XX N-PDB; AAF59012.

XX Novel peptide and its use -  
PT  
v

PS Claim 7; Page 39-40; 75pp; Japanese.

CC The present invention describes a peptide, its complex, derivative or  
CC its polymerizate, where the peptide (I) has a formula of:  
CC alpha1-alpha2-alpha3-alpha4-alpha5-alpha6-alpha7  
CC where alpha1 to alpha7 = amino acid sequence selected from amino acid  
CC sequences ranging from 11-19 amino acids derived from T cell epitopes  
CC derived from cedar (Japanese cedar - Cryptomeria japonica) pollen  
CC allergens. The peptide can be used in an antiasthma/asthma agent.  
CC AAB59011 to AAB59062 and AAB59094 to AAB59121 represent sequences used  
CC in the exemplification of the present invention.

**SQ Sequence 210 AA;**

Query Match	Score	DB	Length
25.0%	178	22	210

Matches 53; Conservative 7; Mismatches 7; Indels 96; Gaps 6;

QY 1 MKVTVAFNQGPNR VFIKRVSNVI IHGRRIDIFASKNFHLQKNT-IGTGR----- 500

QY 1 MKVTVAFNQGPNR VFIKRVSNVI IHGRRIDIFASKNFHLQKNT-IGTGR----- 500

Db 29 MKVTVAFNQFGP-----FASKNFHLQKNTKLTSGKIASCLNYGL 67

QY 51 -----RWKNNRIWL-----QPAKLTGFTLMGRPLWII FSGNMNIKLKMP 90

```
Db 68 VHVANNNDPSGKYE GNIYTKKEAFNVEQFAKLTGFTLMGRA----- 11
```

91 MYIAGYKTFDGRRAEVSYVHNGAKFIRVDGIIAAYQNPASW 133

```
Db 111 -----DPR-----GTAAYQNPASW 125
```

Search completed: April 23, 2003, 18:07:49

Job time : 38 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:54:25 ; Search time 15.5769 Seconds

(without alignments)  
212.987 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406  
Sequence: 1 MKYTVAFNPGFNRVFIKR.....IASRVYDGIITAYQNPASWK 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	24.4	514	1	MPA2_CRYJA
2	74	18.2	375	1	MPA1_CHAOB
3	73	18.0	346	1	MPA1_CUPAR
4	73	18.0	367	1	MPA1_JUNAS
5	72	17.7	234	1	A29B_DROME
6	71	17.5	374	1	SBP_CRYJA
7	70	17.2	152	1	RS18_HUMAN
8	66.5	16.4	196	1	YM07_YEAST
9	62.5	15.4	145	1	YP12_AGRU
10	62.5	15.4	434	1	PEL_LILLO
11	62.5	15.4	1290	1	PIG1_HUMAN
12	62	15.3	173	1	PAPB_ECOLI
13	61.5	15.1	633	1	TOP1_THEMA
14	61	15.0	152	1	RS18_DROME
15	60.5	14.9	110	1	Y12K_SMSV4
16	60	14.8	727	1	PKP1_BOVIN
17	60	14.8	747	1	PKP1_HUMAN
18	59.5	14.7	842	1	PHS2_RAT
19	59	14.5	152	1	RS18_SCHPO
20	59	14.5	234	1	A29B_DROST
21	59	14.5	840	1	MUTS_LACTA
22	59	14.5	862	1	KM2_YEAST
23	58.5	14.4	478	1	YSP3_YEAST
24	58.5	14.4	842	1	PHS2_YEAST
25	58.5	14.4	1291	1	PIG1_BOVIN
26	58	14.3	152	1	RS18_ARATH
27	58	14.3	565	1	HEMA_IADAI
28	58	14.3	646	1	NODO_RH193
29	58	14.3	728	1	PKP1_MOUSE
30	58	14.3	1070	1	YH4_YEAST
31	57.5	14.2	207	1	PIG1_MOUSE
32	57.5	14.2	1290	1	PIG1_RAT
33	57	14.0	379	1	YLEU_DEBOC

34	57	14.0	511	1	UDPGL_DICDI
35	57	14.0	567	1	ODP2_HAEIN
36	56.5	13.9	99	1	IHPA_XYLFPA
37	56.5	13.9	215	1	RAN_BRUMA
38	56.5	13.9	339	1	BMD4_BORBU
39	56.5	13.9	394	1	MEMB_METTR
40	56.5	13.9	842	1	PHS2_HUMAN
41	56.5	13.9	994	1	PROD_DROME
42	56.5	13.9	1163	1	PROD_PEA
43	56	13.8	246	1	RS3_AERPE
44	56	13.8	267	1	RS3_MYCCE
45	56	13.8	341	1	AUR_AQUUPY

## ALIGNMENTS

RESULT 1  
MPA2\_CRYJA STANDARD: PRT; 514 AA.  
AC P4312;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Possible polygalacturonase precursor (BC 3.2.1.15) (PG) (Pectinase)  
DE (Major pollen allergen Cry j 2) (Cry j II).  
OS Crytomeria japonica (Japanese cedar).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cryptomeria.  
OX NCBI\_TaxID=3369;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Pollen;  
RX MEDLINE=95010777; PubMed=7926035;  
RA Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S.,  
RA Usui M., Kurimoto M.;  
RT "Molecular cloning of the second major allergen, Cry j II, from  
RT Japanese cedar pollen.";  
RL FEBS Lett. 353:124-128(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pollen;  
RX MEDLINE=94271186; PubMed=8002972;  
RA Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K.;  
RT "cDNA cloning and expression of Cry j II the second major allergen of  
RT Japanese cedar pollen.";  
RL Biochem. Biophys. Res. Commun. 201:1021-1028(1994).  
RN [3]  
RP SEQUENCE OF 55-64.  
RX MEDLINE=90342988; PubMed=2382797;  
RA Sakauchi M., Inouye S., Tanai M., Ando S., Usui M., Matuhasi T.;  
RT "Identification of the second major allergen of Japanese cedar  
RT pollen.";  
RL Allergy 45:309-312(1990).  
RN [4]  
RP CATALYTIC ACTIVITY: Random hydrolysis of 1,4-alpha-D-  
RN galactosiduronic linkages in pectate and other galacturonans.  
RN -1- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL).  
RN -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES  
RN (POLYGALACTURONASES).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: D37765; BAA07021.1; -  
DR EMBL: D29772; BAA06172.1; -  
DR HSSP: P26505; 1BHE.  
DR Interpro: IPR000743; GH28.  
DR Pfam: PF00295; Glyco\_hydro\_28; 1.

DR PROSITE; PS00502; POLYGALACTURONASE; 1.  
 KM Hydrolyase; Glycosidase; Cell wall; Signal; Zymogen; Fruit ripening;  
 FT Amylolyase; Glycoprotein; Allergen.  
 FT SIGNAL 1 22  
 FT PROPER 23 45  
 FT CHAIN 46 433  
 FT PROPER 434 514  
 FT ACT SITE 278 278  
 FT CARBOHYD 460 460  
 FT CARBOHYD 472 472  
 FT CONFLICT 5 5  
 FT CONFLICT 12 12  
 FT CONFLICT 34 35  
 FT CONFLICT 37 37  
 FT CONFLICT 88 88  
 FT CONFLICT 98 98  
 FT CONFLICT 451 451  
 FT CONFLICT 454 454  
 FT CONFLICT 504 504  
 FT CONFLICT 507 507  
 SQ SEQUENCE 514 AA; 56645 MW; 624611C3FA8D6302 CRC64;

Query Match 24.4%; Score 99; DB 1; Length 514;  
 Best Local Similarity 100.0%; Pred. No. 0.00031;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

31 IDIPASKNFHLOKNTIGT 49  
 DB 236 IDIPASKNFHLOKNTIGT 254

RESULT 2  
 MPAL CHAOS STANDARD; PRT; 375 AA.  
 ID MPAL CHAOS STANDARD; PRT; 375 AA.  
 AC 096385;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Major pollen allergen Cha o 1 precursor.  
 OS Chamaecyparis obtusa (Japanese cypress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae;  
 OC Chamaecyparis.  
 NCBI\_TaxID=13415;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE-Pollen;  
 RX MEDLINE=96255194; PubMed=8676896;  
 RA Suzuki M., Komiyama N., Itoh M., Sone T., Kuno K., Takagi I.,  
 RA Ohta N.;  
 RT "Purification, characterization and molecular cloning of Cha o 1, a  
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen.";  
 RL Mol. Immunol. 33:451-460(1996).  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; DA5404; BAA08246.1; -;  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PRO0807; AMBALLERGEN.  
 KW Allergen; Glycoprotein; Signal.  
 FT CHAIN 1 21  
 FT SIGNAL 22 375  
 FT CARBOHYD 110 110  
 FT CARBOHYD 148 148  
 MAJOR POLLEN ALLERGEN CHA O 1.  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 178 178  
 FT CARBOHYD 293 293  
 FT CARBOHYD 352 352  
 SQ SEQUENCE 375 AA; 40258 MW; 81CD91DF06EDBBF CRC64;

Query Match 18.2%; Score 74; DB 1; Length 375;  
 Best Local Similarity 59.3%; Pred. No. 0.21;  
 Matches 16; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

1 MKVTVAFNQFGNRRVFIKRVSVIIH 27  
 DB 234 MKVTVAFNQFGNRRVFIKRVSVIIH 260

RESULT 3  
 MPAL CUPAR STANDARD; PRT; 346 AA.  
 ID MPAL CUPAR STANDARD; PRT; 346 AA.  
 AC 09SCG9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Major pollen allergen Cup a 1.  
 OS Cupressus arizonica.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Cupressus.  
 NCBI\_TaxID=49011;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20571526; PubMed=11122214;  
 RA Acetunio E., Del Pozo V., Minguez A., Arrieta I., Cortegano I.,  
 RA Cardaba B., Gallardo S., Rojo M., Palomino P., Lahoz C.;  
 RT "Molecular cloning of major allergen from Cupressus arizonica pollen:  
 RT Cup a 1.";  
 RT Clin. Exp. Allergy 30:1750-1758 (2000).  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; AJ243570; CAB62551.1; -;  
 DR InterPro; IPR002022; Amb allergen.  
 DR Pfam; PF00544; pec\_lyase; 1.  
 DR PRINTS; PRO0807; AMBALLERGEN.  
 KW Allergen; Glycoprotein.  
 FT CARBOHYD 127 127  
 FT CARBOHYD 157 157  
 FT CARBOHYD 272 272  
 SQ SEQUENCE 346 AA; 37589 MW; F1281DDDA1D5DFD0 CRC64;

Query Match 18.0%; Score 73; DB 1; Length 346;  
 Best Local Similarity 55.6%; Pred. No. 0.25;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

1 MKVTVAFNQFGNRRVFIKRVSVIIH 27  
 DB 213 MKVTVAFNQFGNRRVFIKRVSVIIH 239

RESULT 4  
 MPAL JUNAS STANDARD; PRT; 367 AA.  
 ID MPAL JUNAS STANDARD; PRT; 367 AA.  
 AC P81294; Q9ZNT7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Major pollen allergen Jun a 1 precursor.  
 OS Juniperus ashei (Ozark white cedar).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Cupressaceae; Juniperus.  
 OX NCB1\_TaxID=13101;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 31-38; 42-50; 58-80; 88-94;  
 RC 117-124; 134-140; 160-164; 236-263 AND 322-325.  
 RA TISSUE=Pollen;  
 RX MEDLINE=99414163; PubMed=10482836;  
 RA Midoro-Horinai T.M., Goldblum R.M., Kurosky A., Wood T.G.,  
 RA Brooks E.G.;  
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major  
 RT allergen, Jun a 1.",  
 RL J. Allergy Clin. Immunol. 104:613-617(1999).  
 RN [2]  
 RP SEQUENCE OF 22-50.  
 RC TISSUE=Pollen;  
 RX MEDLINE=99414162; PubMed=10482835;  
 RA Midoro-Horinai T., Goldblum R.M., Kurosky A., Goetz D.W.,  
 RA Brooks E.G.;  
 RT "Isolation and characterization of the mountain cedar (Juniperus  
 RT ashei) pollen major allergen, Jun a 1.",  
 RL J. Allergy Clin. Immunol. 104:608-612(1999).  
 CC -1- DISEASE: THIS PROTEIN CAUSES SEVERAL SEASONAL ALLERGIC RHINITIS  
 CC IN NORTH AMERICA.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A I/AMB A II/CRY J I SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF106663; AAD03609.1; -;  
 CC EMBL; AF106662; AAD03608.1; -;  
 CC InterPro: IPR002022; Amb\_allergen.  
 CC Pfam: PF00544; pec\_lyase; 1.  
 CC PRINTS; PR00807; AMBALLERGEN.  
 CC KMW Allergen; Glycoprotein; Signal.  
 CC FT SIGNAL 21  
 CC FT CHAIN 1 21 MAJOR POLLEN ALLERGEN JUN A 1.  
 CC FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SQ SEQUENCE 367 AA; 39824 MW; FC9B81E675662E49 CRC64;  
 CC  
 CC Query Match 18.0%; Score 73; DB 1; Length 367;  
 CC Best Local Similarity 55.6%; Pred. No. 0.27;  
 CC Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 CC  
 CC QY 1 MKTVAPNGGKRRVPIKRVNVIIH 27  
 CC DB 234 MKTVAPNGFGPNAGQMPRRARYGLVH 260  
 CC  
 CC RESULT 5  
 CC A29B DROME STANDARD; PRT; 234 AA.  
 CC AC Q46197; Q9TV73; Q9TW05; Q9TW07; Q9U976; Q9U977; Q9U978;  
 CC AC Q9U979; Q9V305;  
 CC DT 16-OCT-2001 (Rel. 40, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Accessory gland protein Acp29AB precursor.  
 CC GN ACP29AB OR CG1797.  
 CC OS Drosophila melanogaster (Fruit fly).  
 CC OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 CC OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 CC OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 CC OX NCB1\_TaxID=7227;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

RC STRAIN=Canton-S; TISSUE=Male accessory gland;  
 RX MEDLINE=98135120; PubMed=9474779;  
 RA Wolfner M.F., Harada H.A., Berman M.J., Stelick T.J., Kraus K.W.,  
 RA Kalb J.M., Lung Y.O., Neuhau D.M., Park M., Tram U.K.;  
 RT "New genes for male accessory gland proteins in Drosophila  
 RT melanogaster.",  
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Various strains;  
 RX MEDLINE=99282496; PubMed=10353898;  
 RA Aguade M.;  
 RT "Positive selection drives the evolution of the Acp29AB accessory  
 RT gland protein in Drosophila.",  
 RL Genetics 152:543-551(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Various strains;  
 RX MEDLINE=20556153; PubMed=11102381;  
 RA Begun D.J., Whitely P., Todd B.L., Waldrip-Dail H.M., Clark A.G.;  
 RT "Molecular population genetics of male accessory gland proteins in  
 RT Drosophila.",  
 RL Genetics 156:1879-1888(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratunga C., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abul J.F., Agbayani A., An H.-D., Andrews-Plamkoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter L.,  
 RA Burlingame K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,  
 RA de Paulis B., Delcher A., Deng Z., Mayhew A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Evans G.L., Fouts D.E., Garg N.S., Gerhart M., Glasser K.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matrei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshireff A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.E., Smith T.,  
 RA Stryker R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.",  
 RL Science 287:2185-2195(2000).  
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
 CC MATED MALE FLIES.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: MAIN CELLS OF THE ACCESSORY GLAND AND IN  
 CC SEMINAL FLUID.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; U85758; AAB96382.1; -  
 DR EMBL; AJ240513; CAB53187.1; -  
 DR EMBL; AJ240514; CAB53188.1; -  
 DR EMBL; AJ240515; CAB53189.1; -  
 DR EMBL; AJ240516; CAB53190.1; -  
 DR EMBL; AJ240517; CAB53191.1; -  
 DR EMBL; AJ240518; CAB53192.1; -  
 DR EMBL; AJ240519; CAB53193.1; -  
 DR EMBL; AJ240520; CAB53194.1; -  
 DR EMBL; AJ240521; CAB53195.1; -  
 DR EMBL; AJ240522; CAB53196.1; -  
 DR EMBL; AJ240523; CAB53197.1; -  
 DR EMBL; AJ240524; CAB53198.1; -  
 DR EMBL; AJ240525; CAB53199.1; -  
 DR EMBL; AJ240526; CAB53200.1; -  
 DR EMBL; AJ240527; CAB53201.1; -  
 DR EMBL; AJ240528; CAB53202.1; -  
 DR EMBL; AJ240529; CAB53203.1; -  
 DR EMBL; AJ240530; CAB53204.1; -  
 DR EMBL; AJ240531; CAB53205.1; -  
 DR EMBL; AJ240532; CAB53206.1; -  
 DR EMBL; AJ240533; CAB53207.1; -  
 DR EMBL; AJ240534; CAB53208.1; -  
 DR EMBL; AJ240535; CAB53209.1; -  
 DR EMBL; AJ240536; CAB53210.1; -  
 DR EMBL; AJ240537; CAB53211.1; -  
 DR EMBL; AJ240538; CAB53212.1; -  
 DR EMBL; AJ240539; CAB53213.1; -  
 DR EMBL; AJ240540; CAB53214.1; -  
 DR EMBL; AJ240541; CAB53215.1; -  
 DR EMBL; AJ240542; CAB53216.1; -  
 DR EMBL; AJ240543; CAB53217.1; -  
 DR EMBL; AJ240544; CAB53218.1; -  
 DR EMBL; AJ240545; CAB53219.1; -  
 DR EMBL; AJ240546; CAB53220.1; -  
 DR EMBL; AJ240547; CAB53221.1; -  
 DR EMBL; AJ240548; CAB53222.1; -  
 DR EMBL; AJ240549; CAB53223.1; -  
 DR EMBL; AJ240550; CAB53224.1; -  
 DR EMBL; AJ240551; CAB53225.1; -  
 DR EMBL; AY010527; AAG32559.1; -  
 DR EMBL; AY010528; AAG32560.1; -  
 DR EMBL; AY010529; AAG32561.1; -  
 DR EMBL; AY010530; AAG32562.1; -  
 DR EMBL; AY010531; AAG32563.1; -  
 DR EMBL; AY010532; AAG32564.1; -  
 DR EMBL; AY010533; AAG32565.1; -  
 DR EMBL; AY010534; AAG32566.1; -  
 DR EMBL; AY010535; AAG32567.1; -  
 DR EMBL; AY010536; AAG32568.1; -  
 DR EMBL; AY010537; AAG32569.1; -  
 DR EMBL; AY010538; AAG32570.1; -  
 DR EMBL; AY010539; AAG32571.1; -  
 DR EMBL; AY010540; AAG32572.1; -  
 DR EMBL; AY010541; AAG32573.1; -  
 DR EMBL; AY010542; AAG32574.1; -  
 DR EMBL; AY010543; AAG32575.1; -  
 DR EMBL; AEO03621; AAF52655.1; -  
 DR FLYBase; FBgn0015583; ACP29AB.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR Pfam; PF00059; Lectin\_C\_1.  
 DR PROSITE; PS00615; C-TYPE LECTIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50041; C-TYPE LECTIN\_2; 1.  
 KW Behavior; Lectin; Signal; Polymorphism.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 22 234 ACCESSORY GLAND PROTEIN ACP29AB.

FT DOMAIN 137 228 C-TYPE LECTIN (LONG FORM).  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 29 29 K -> N (IN STRAINS MA7, WS16 AND WS26).  
 FT VARIANT 59 59 K -> R (IN STRAINS LA1, LA3, LA4, LA5, LA14, LA15, LA16, LA21, LA25, LA34, LA35, MA18, MA19, MA21, MA45, MA48, MA52, MA57, MA20, MA21, MA45, MA48, MA52, MA57, MA20, ZIM29, ZIM30, ZIM56, WS6 AND WS19).  
 FT VARIANT 105 105 A -> S (IN STRAINS LA16 AND LA35).  
 FT VARIANT 113 113 R -> L (IN STRAINS LA14, LA14, LA16, LA35, MA28, MA15B, MA34A, MA36A, MA37A, MA47A, MA52B, MA58B, MA80B, MA5, MA7, MA13, MA18, MA20, MA21, MA45, MA48, MA50, MA51, MA67, WS1, WS6, WS9, WS12, WS16, WS26, WS47, WS56, ZIM2, ZIM26, ZIM29, ZIM30, ZIM32, ZIM37, ZIM42 AND ZIM56).  
 FT VARIANT 153 153 K -> M (IN STRAINS BERKELEY, LA13, LA14, LA15, LA16, MA18, MA8B, MA40B, MA52B, MA5, MA21, MA45, MA52, MA67, ZIM29, ZIM30, ZIM42 AND ZIM56).  
 FT VARIANT 214 214 E -> D (IN STRAIN MA7).  
 FT SEQUENCE 234 AA; 27173 MW; 8954CD3215480F3B CRC64;  
 SQ  
 Query Match 17.7%; Score 72; DB: 1; Length 234;  
 Best Local Similarity 25.9%; Fred. NO. 0.22; 31; Indels 10; Gaps 3;  
 Matches 21; Conservative 19; Mismatches 19;  
 QY 9 QFGPNRVRFTKRSNV---IIHGRRIDFASKNPHLQNTTGT--GRRISIKLTSGKIA 62  
 DB 88 QLOPLKTIWRHNSNKAANNIKRRFEKVSRRHIENLMTQTFEAYVTCRKNGHILA 147  
 QY 63 S---RRVDGIITAYONPASW 79  
 DB 148 NIDKEKELDGIILAPNNNSYW 168  
 ID SBP CRYOA STANDARD; PRT; 374 AA.  
 AC P18632;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Sugi baatic protein precursor (SBP) (Major allergen Cry j 1) (Cry j I).  
 OS Cryptomeria japonica (Japanese cedar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Cupressales; Cupressaceae; Cryptomeria.  
 OC Spermatophyta; Coniferales; Coniferales; Cupressaceae; Cryptomeria.  
 OX NCBI\_Taxid=3369;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Pollen;  
 RX MEDLINE=94183234; PubMed=8135802;  
 RA Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K., Kuno K.;  
 RT "Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese cedar pollen.";  
 RT Biochem. Biophys. Res. Commun. 199;619-625 (1994).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pollen;  
 RA Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M.;  
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 22-41.  
 RC TISSUE=Pollen;  
 RX MEDLINE=89031257; PubMed=3181436;  
 RA Tanaka M., Ando S., Ueki M., Kurimoto M., Sakaguchi M., Inoue S., Marubasi T.;  
 RT "N-terminal amino acid sequence of a major allergen of Japanese cedar pollen (Cry j I).";  
 RT FEBS Lett. 239;329-332 (1988).



[4]  
 RN CARBOHYDRATES.  
 RP TISSUE=Polien;  
 RX MEDLINE=95003748; PubMed=7920021;  
 RA Hijikata A., Matsunoto I., Kojima K., Ogawa H.;  
 RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar  
 (Cryptomeria japonica) pollen allergen, Cry j1.";  
 RL Int. Arch. Allergy Immunol. 105:198-202(1994).  
 RN [5]  
 RP STRUCTURE OF CARBOHYDRATES.  
 RX TISSUE=Polien;  
 MEDLINE=95332249; PubMed=7608114;  
 RA Hino K., Yamamoto S., Sano O., Taniguchi Y., Kohno K., Usui M.,  
 RT Fukuda S., Hanzawa H., Hanyama H., Kurimoto M.;  
 RT Carbohydrate structures of the glycoprotein allergen Cry j 1 from  
 Japanese cedar (Cryptomeria japonica) pollen.";  
 RL J. Biochem. 117:289-295(1995).  
 CC -1- PPM: CONTAINS FUCCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.  
 CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR  
 CC POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN.  
 CC -1- MISCELLANEOUS: THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM  
 CC B DIFFERS IN SIX POSITIONS.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC AMB A II/CRY J I SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D26544; BAA05542.1; -;  
 DR EMBL; D26545; BAA05543.1; -;  
 DR EMBL; D34639; BAA07020.1; -;  
 DR PIR; A44773; A44773.  
 DR GICOSuiteDB: p18632; -;  
 DR InterPro: IPR002022; Amb-allergen.  
 DR Pfam: PF00544; pec\_lyase; 1.  
 DR PRINTS; PR00807; AMBALLERGEN.  
 KW Allergen; Glycoprotein; Multigene family; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21  
 FT CARBOHYD 158 374  
 FT CARBOHYD 191 158  
 FT CARBOHYD 191 158  
 FT CARBOHYD 293 354  
 FT CARBOHYD 354 354  
 FT VARIANT 12 12  
 FT VARIANT 143 143  
 FT VARIANT 202 202  
 FT VARIANT 221 221  
 FT VARIANT 358 358  
 FT VARIANT 361 361  
 FT VARIANT 361 361  
 SQ SEQUENCE 374 AA, 40645 MW, 74AB25950248556F CRC64;  
 Query Match 17.5%; Score 71; DB 1; Length 374;  
 Best Local Similarity 55.6%; Pred. No. 0.48;  
 Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

DT 15-UN-2002 (rel. 41, last annotation update)  
 DE 40S ribosomal protein S18 (KE-3) (KE3).  
 GN RPS18.  
 OS Homo sapiens (Human).  
 OS Mus musculus (Mouse).  
 OS Rattus norvegicus (Rat), and  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606, 10090, 10116, 9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human; TISSUE=Placenta;  
 RX MEDLINE=93181276; PubMed=8441687;  
 RA Chassin D., Ballet D., Koman A.;  
 RT "The human homolog of ribosomal protein S18.";  
 RL Nucleic Acids Res. 21:745-745(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Human;  
 RA Tubby B.;  
 RT Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Rat; STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RX MEDLINE=91337062; PubMed=1872840;  
 RA Chan Y.-L., Paz V., Wool I.G.;  
 RT "The primary structure of rat ribosomal protein S18.";  
 RL Biochem. Biophys. Res. Commun. 178:1212-1218(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Mouse;  
 RX MEDLINE=92182530; PubMed=1543907;  
 RA Macmurray A.J., Shin H.S.;  
 RT "The murine MHC encodes a mammalian homolog of bacterial ribosomal  
 RT protein S13.";  
 RL Mamm. Genome 2:87-95(1992).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Mouse; STRAIN=129/SvJ;  
 RA Rowen L., Qin S., Madan A., Loretz C., James R., Dors M., Mix L.,  
 RA Hall J., Lasky S., Hood L.;  
 RT "Sequence of the mouse major histocompatibility locus class II  
 RT region.";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Pig;  
 RA Kimura M., Kawakami K., Suzuki H., Hamasima N.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE  
 CC INITIATION OF TRANSLATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X57529; CAA40750.1; -;  
 DR EMBL; AL031228; CAA20231.1; -;  
 DR EMBL; X69150; CAB56794.1; -;  
 DR EMBL; W76763; AAA16795.1; -;  
 DR EMBL; W76762; AAA16796.1; -;  
 DR EMBL; AF100956; AAC69898.1; -;  
 DR EMBL; AF110520; AAC97978.1; -;  
 DR EMBL; AB000911; BAA19211.1; -;  
 DR PIR; JH0419; R3RT18.  
 DR PIR; S30393; S30393.  
 DR Genew; HGNC:10401; RPS18.

MM; 180473; -.  
 DR MGD; MGI:98146; Rps18.  
 DR InterPro; IPR001892; Ribosomal\_S13.  
 DR Pfam; PF00416; Ribosomal\_S13; 1.  
 DR ProDom; PD001363; Ribosomal\_S13; 1.  
 DR PROSITE; PS00646; RIBOSOMAL\_S13; 1.  
 KW Ribosomal protein.  
 FT CONFLICT 104 104 D -> H (IN REF. 4; AAL6796).  
 SQ SEQUENCE 152 AA; 17719 MW; 4DAF0662C3F37F22 CRC64;  
 Query Match 17.2%; Score 70; DB 1; Length 152;  
 Best Local Similarity 30.4%; Pred. No. 0.23;  
 Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

OY 12 PNRKPRIVKRVSVITHGRRIDIFASKNPHLOKNTIGTGR-----ISLKTSGKI 61  
 DB 6 PEKFGHILRVLTNTIDGRKIAFALTAK-----GVGRYAHVVLKRAKIDLTFRAGEL 59  
 OY 62 ASRVGDIIAAYONPASKW 80  
 DB 60 TEDEVYRITTMQNPQYK 78

RESULT 8  
 YMO7\_YEAST STANDARD; PRT; 196 AA.  
 ID\_YMO7\_YEAST  
 AC 004487;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Putative succinate dehydrogenase cytochrome B subunit, mitochondrial precursor.  
 GN YMR18C OR YMR718.17C.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes;  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RA Hunt S., Bowman S., Barrell B.G., Rajandream M.A.;  
 RL Submitted (May-1995) to the EMBL/GenBank/DDJ databases.  
 CC -1- FUNCTION: MONO-HEME CYTOCHROME B. INVOLVED IN SYSTEM II OF THE  
 CC MITOCHONDRIAL ELECTRON TRANSPORT CHAIN WHICH IS RESPONSIBLE FOR  
 CC TRANSFERRING ELECTRONS FROM SUCCINATE TO UBIQUINONE (COENZYME Q)  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B560 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z49702; CAAB9756.1; -.  
 CC SGD; S0004724; YMR18C.  
 DR InterPro; IPR000701; Sdh\_cyt.  
 DR Pfam; PF01127; Sdh\_cyt; 1.  
 DR PROSITE; PS01000; SDH\_CYT\_1; 1.  
 DR PROSITE; PS01001; SDH\_CYT\_2; 1.  
 KW Hypothetical protein; Tricarboxylic acid cycle; Electron transport;  
 KW Heme; Transmembrane; Mitochondrion; Transit peptide.  
 FT TRANSIT 1 2 MITOCHONDRION (POTENTIAL).  
 FT CHAIN 1 196 PUTATIVE SUCCINATE DEHYDROGENASE  
 FT CYTOCHROME B SUBUNIT.  
 FT TRANSMEM 99 119 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 SQ SEQUENCE 196 AA; 22309 MW; 4141398B9B2B057 CRC64;

Query Match 16.4%; Score 66.5; DB 1; Length 196;  
 Best Local Similarity 29.7%; Pred. No. 0.81;  
 Matches 27; Conservative 16; Mismatches 33; Indels 15; Gaps 7;

OY 1 MKVNV--AFNPG--PNRNVFIKRV-----NVIIHGRRIDIFASKNPH--LOKNTIGT 48  
 DB 1 MKATIQRTSVGVGPRASVFPRISTPILHNTYISNG-RMDLF-SKEFHNGRVSKSLWS 58

OY 49 GRISLKTSGKIASRRVDGIIAAYONPASKW 79  
 DB 59 SNKEELVSGR-KRPISPHLTYVEPEMSW 88

RESULT 9  
 YP12\_AGRU STANDARD; PRT; 145 AA.  
 ID\_YP12\_AGRU  
 AC 044433;  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 15.8 kDa protein in pin2.3'region (ORF2).  
 DE Agrobacterium tumefaciens.  
 OS Plasmid pTiA6.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OC NCBI\_TaxID=358;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89213933; PubMed=2708311;  
 RA Kanemoto R.H., Powell A.T., Akiyoshi D.B., Regier D.A.,  
 RA Kerstetter R.A., Neeter E.W., Hawes M.C., Gordon M.P.;  
 RT Nucleotide sequence and analysis of the plant-inducible locus pinP  
 RT from Agrobacterium tumefaciens.  
 RL J. Bacteriol. 171:2506-2512 (1989).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M19352; AAA82504.1; -.  
 KW Hypothetical protein; Plasmid.  
 KM SQ SEQUENCE 145 AA; 15829 MW; 95A3D3D8B560AA6 CRC64;  
 SQ

Query Match 15.4%; Score 62.5; DB 1; Length 145;  
 Best Local Similarity 33.3%; Pred. No. 1.7;  
 Matches 17; Conservative 11; Mismatches 18; Indels 5; Gaps 2;

OY 18 IKRVSNVITH-GRRIDIFASKNPHLOKNTIGTGRKRSIKLTSGKIASRRVD 67  
 DB 1 MKRISTIVGVPLATPYAADNTH---TLGTLSIELALTKAPVAVTVTD 47

RESULT 10  
 PEL\_LILLO STANDARD; PRT; 434 AA.  
 ID\_PEL\_LILLO  
 AC P40973;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Pectate lyase precursor (BC 4.2.2.2).  
 DE Liliium longilorum (Trumpet Lily).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;  
 OC Liliium.  
 OC NCBI\_TaxID=4690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nellie white; TISSUE=Pollen;  
 RA Kim S.R., Finkel D.J., An G.;

RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: Eliminative cleavage of pectate to give  
 CC oligosaccharides with 4-deoxy-alpha-D-gluc-4-enurosyl groups at  
 CC their non-reducing ends.  
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z17328; CAA78976.1; -;  
 DR EMBL: L18911; AAA33398.1; -;  
 DR PIR: S29612; S29612.  
 DR InterPro: IPR002023; Amb allergen.  
 DR Pfam: PF00544; Pec\_lyase\_1.  
 DR PRINTS: PR00807; AMBALLERGEN.  
 DR Lyase; Signal.  
 KM SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 434 PEPTATE LYASE.  
 FT ACT SITE 312 312 POTENTIAL.  
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;  
 Query Match 15.4%; Score 62.5; DB 1; Length 434;  
 Best Local Similarity 28.4%; Pred. No. 5.8; Indels 19; Gaps 4;  
 Matches 25; Conservative 11; Mismatches 33;  
 QY 1 MKYTVANFGNCPN-----RRVFIRVSNVTHIGRRIDIPASKNFHLQKNTIGTGR- 51  
 DB 296 MQYTVANFHNHGRGLVGRMPRCRYGFVHVNDYTHIMVAVGSGR-----PTISGKNY 351  
 QY 52 ISLKTSGKIASRRVDIIAAYQNPASW 79  
 DB 352 IAPHIEAKEVTKR-----DYAEPAEW 373  
 RESULT 11  
 P1GI1 HUMAN STANDARD; PRT; 1290 AA.  
 AC P19174;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma 1  
 DE (EC 3.1.4.11) (PLC-gamma-1) (Phospholipase C-gamma-1) (PLC-11)  
 DE (PLC-148).  
 GN PLCG1 OR PGL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metazoa; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID:9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=vein, and Brain;  
 RX MEDLINE=9035593; PubMed=2167438;  
 RA Burgess W.H., Dione C.A., Kaplow J.M., Mudd R., Friesel R.,  
 RA Zilberstein A., Schlesinger J., Jaye M.;  
 RT "Characterization and cDNA cloning of phospholipase C-gamma, a major  
 RT substrate for heparin-binding growth factor I (acidic fibroblast  
 RT growth factor)-activated tyrosine kinase.";  
 RT Mol. Cell. Biol. 10:4770-4777(1990).  
 RU [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21638749; PubMed=11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,

RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Cobby N.R.,  
 RA Coulson A., Coville G.D., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Grahame D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle B., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levasseur M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McKay K., McMurran A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillips B.J.C.T., Pratchinham S.R., Plumb R.W., Ramsey H.,  
 RA Rice C.W., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,  
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RA "The DNA sequence and comparative analysis of human chromosome 20.";  
 RA Nature 414:865-871(2001).  
 [3]  
 RX STRUCTURE BY NMR OF SH3 DOMAIN.  
 RX MEDLINE=93208890; PubMed=7681365;  
 RA Kohda D., Hatanaka H., Odaka M., Mandayan V., Ullrich A.,  
 RA Schlesinger J., Inagaki F.;  
 RT "Solution structure of the SH3 domain of phospholipase C-gamma.";  
 RT Cell 72:953-960(1993).  
 CC -1- FUNCTION: PLC-GAMMA IS A MAJOR SUBSTRATE FOR HEPARIN-BINDING  
 CC GROWTH FACTOR 1 (ACIDIC FIBROBLAST GROWTH FACTOR)-ACTIVATED  
 CC TYROSINE KINASE.  
 CC -1- CATALYTIC ACTIVITY: 1-phosphatidy-1-D-myo-inositol 4,5-  
 CC bisphosphate + H(2)O = D-myo-inositol 1,4,5-trisphosphate +  
 CC diacylglycerol.  
 CC -1- COFACTOR: Calcium.  
 CC -1- PTM: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-GAMMA 1 AND 2  
 CC INVOLVES THEIR PHOSPHORYLATION BY TYROSINE KINASES IN RESPONSE  
 CC TO LIGATION OF A VARIETY OF GROWTH FACTOR RECEPTORS AND IMMUNE  
 CC SYSTEM RECEPTORS.  
 CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS  
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.  
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 PH DOMAINS. THE SECOND ONE IS IN TWO  
 CC PARTS.  
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EF-HAND CALCIUM-BINDING DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M34667; AAA36452.1; -;  
 DR EMBL: A0022394; CAA18537.1; -;  
 DR PIR: A36466; A36466.  
 DR PDB: 2HSF; 31-AUG-94.  
 DR PDB: 1HSQ; 31-AUG-94.  
 DR Genew: HGNC:9065; PLCG1.  
 DR MIM: 172420;  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR002048; EF-hand.  
 DR InterPro: IPR001849; PH.  
 DR InterPro: IPR001192; PI\_PLC.  
 DR InterPro: IPR000909; PI\_PLC\_Xdom.  
 DR InterPro: IPR001711; PI\_PLC\_Y.  
 DR InterPro: IPR000980; SH2.  
 DR InterPro: IPR001452; SH3.

DR Pfam; PF00017; SH2; 2.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00169; PH; 2.  
 DR Pfam; PF00387; PI-PLC-Y; 1.  
 DR Pfam; PF00388; PI-PLC-X; 1.  
 DR PRINTS; PR00390; PHPLIPASEC.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR Pfam; PD000066; SH3; 1.  
 DR Pfam; PD000093; SH2; 2.  
 DR Pfam; PD001202; PI-PLC-Y; 1.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00233; PH; 2.  
 DR SMART; SM00148; PLCYC; 1.  
 DR SMART; SM00149; PLCYC; 1.  
 DR SMART; SM00252; SH2; 2.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00018; EF\_HAND; 1.  
 DR PROSITE; PS50001; SH2; 2.  
 DR PROSITE; PS50002; SH3; 1.  
 DR PROSITE; PS50003; PH\_DOMAIN; 2.  
 DR PROSITE; PS50007; PIPLC\_X\_DOMAIN; 1.  
 DR PROSITE; PS50008; PIPLC\_Y\_DOMAIN; 1.  
 KW Hydrolase; Lipid degradation; Transducer; SH2 domain; SH3 domain;  
 Repeat; Calcium-binding; Phosphorylation; 3d-structure; Polymorphism.  
 FT DOMAIN 27 142 EF\_HAND (POTENTIAL).  
 FT CA\_BIND 165 176 DOMAIN X.  
 FT DOMAIN 320 464 PH 2 (FIRST PART).  
 FT DOMAIN 489 523 SH2 1.  
 FT DOMAIN 550 657 SH2 2.  
 FT DOMAIN 668 756 SH3.  
 FT DOMAIN 791 851 PH 2 (SECOND PART).  
 FT DOMAIN 953 1070 DOMAIN Y.  
 FT ACT\_SITE 1075 1177 C2 DOMAIN.  
 FT ACT\_SITE 335 335 BY SIMILARITY.  
 FT ACT\_SITE 380 380 BY SIMILARITY.  
 FT MOD\_RES 771 771 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 783 783 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 1253 1253 PHOSPHORYLATION (BY SIMILARITY).  
 FT VARIANT 813 813 I->T (IN DBSNP:753811).  
 FT VARIANT 813 813 /FTID=VAR\_011908.  
 SQ SEQUENCE 1290 AA; 148531 MW; AE05ABE2A18EDDAC CRC64;  
 Query Match 15.4%; Score 62.5; DB 1; Length 1290;  
 Best Local Similarity 28.1%; Pred. No. 19;  
 Matches 18; Conservative 9; Mismatches 34; Indels 3; Gaps 1;  
 QY 5 VAFNQGPNRRVFIKRVSVIIHGRRIIDIFASKNFHLQ--KNITGTRRISLKLTSKGI 61  
 Db 883 IAIIRPEGKNNRLVFSISMAVSHWMLDVAADSOEELQOWVKIKREVAQTADARLTGKI 942  
 QY 62 ASRR 65  
 Db 943 MERR 946  
 RESULT 12  
 ID PAPE\_ECOLI STANDARD; PRT; 173 AA.  
 AC P08407;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Fimbrial protein pape precursor.  
 GN PAPE  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_Taxid=562;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=J96;  
 RX MEDLINE=86149403; PubMed=2869489;  
 RA Lindberg F., Lund B., Normark S.;  
 RT "Gene products specifying adhesion of uropathogenic Escherichia coli  
 RT are minor components of pili";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:1891-1895(1986).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J96;  
 RX MEDLINE=88169520; PubMed=2895103;  
 RA Lund B., Lindberg F., Normark S.;  
 RT "Structure and antigenic properties of the tip-located P pilus  
 RT proteins of uropathogenic Escherichia coli";  
 RL J. Bacteriol. 170:1887-1894(1988).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J96;  
 RX MEDLINE=93023852; PubMed=1357526;  
 RA Marklund B.-I., Tennent J.M., Garcia E., Hamers A., Baga M.,  
 RA Lindberg F., Gaastria W., Normark S.;  
 RT "Horizontal gene transfer of the Escherichia coli pap and prs pili  
 RT operons as a mechanism for the development of tissue-specific  
 RT adhesive properties";  
 RL Mol. Microbiol. 6:2225-2242(1992).  
 RL [4]  
 CC -1- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING  
 CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5  
 CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO  
 CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.  
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE MINOR COMPONENTS OF PILI.  
 CC PILI WITH A DEFECTIVE PAPE GENE WILL HAVE LOW ADHESIVE CAPACITY  
 CC OR NONE, HOWEVER, THE BINDING PROPERTY OF THE WHOLE CELL WILL NOT  
 CC BE AFFECTED.  
 CC -1- SUBCELLULAR LOCATION: SECRETED; TIP OF THE PILI.  
 CC -1- DISEASE: STRAINS OF E. COLI THAT CAUSE INFECTION OF THE HUMAN  
 CC URINARY TRACT PRODUCE PAP-PILI WHICH ARE HAIR-LIKE APPENDAGES  
 CC CONSISTING OF ABOUT 1000 HELICALLY ARRANGED SUBUNITS OF THE  
 CC PROTEIN PAPE. THESE PILI MEDIATE BINDING TO DIGALACTOSIDE-  
 CC CONTAINING GLYCOPOLYIDS PRESENT ON THE EPITHELIAL CELLS WHICH  
 CC LINE THE URINARY TRACT.  
 CC [5]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X61238; CAA43556.1; -  
 CC EMBL; M33239; AAA4280.1; -  
 CC EMBL; M20146; AAA4288.1; -  
 CC EMBL; X61239; CAA43568.1; -  
 CC PIR; A25134; YOECEP.  
 CC InterPro; IPR000259; Fimbrial.  
 CC InterPro; IPR004086; FimbrialPape.  
 CC Pfam; PF00419; Fimbrial; 1.  
 CC PRINTS; PR01555; FIMBRIALPAPE.  
 KW Fimbria; Cell adhesion; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 173 FIMBRIAL PROTEIN PAPE.  
 SQ SEQUENCE 173 AA; 18569 MW; E27577D09C46A863 CRC64;  
 Query Match 15.3%; Score 62; DB 1; Length 173;  
 Best Local Similarity 27.0%; Pred. No. 24;  
 Matches 20; Conservative 11; Mismatches 35; Indels 8; Gaps 3;  
 QY 1 MKKTVV-AFNQGPNNRRVFIKRVSVIIHGRRIIDIFASKNFHLQKNTIGTRRISLKLTSK 59  
 Db 79 MKVITITATNTY--NNALIVQNTSITSSDGLLVLYVNS-----NAGNIGTATITGTFTPTG 131  
 QY 60 KIASRVVDGIITAY 73

DB 132 KITGNNAKDTISLH 145

RESULT 13

TOPI1 THEME STANDARD; PRT; 633 AA.

AC P46759; 16-OCT-2001 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)

GN (Unwinding enzyme) (Swivelase).

OS TOPA OR TM0258.

OC Thermotoga maritima.

OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;

OC Thermotogaceae; Thermotoga.

OX NCBI\_TaxID=2336;

FN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=96138548; PubMed=8547314;

RA Bouchier de la Tour C., Kalloum H., Portemer C., Confalonieri F., Huber R., Duguet M.,

RT "Cloning and sequencing of the gene coding for topoisomerase I from the extremely thermophilic eubacterium, Thermotoga maritima.";

RL Biochim. Biophys. Acta 1264:279-283 (1995).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=99287316; PubMed=10360571;

RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A., McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M., Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D., Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O., Salzberg S.L., Smith H.O., Venter J.C., Frazer C.M., White O.,

RT "Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima.";

RL Nature 399:323-326 (1999).

CC -1- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.

CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded DNA, followed by passage and rejoining.

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS AT ONE END OF THE ENZYME-SEVERED DNA STRAND.

CC -1- SIMILARITY: BELONGS TO PROKARYOTIC TYPE I/III TOPOISOMERASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: U27841; AAA68949.1; -

CC EMBL: AE001708; AAD35346.1; -

CC HSSP: P06612; 1ECL.

CC TIGR: TM0258; -

CC InterPro: IPR002936; DMAPrim\_toprim.

CC InterPro: IPR003601; DNATopi\_ATP\_bind.

CC InterPro: IPR003602; DNATopi\_DNA\_bind.

CC InterPro: IPR00380; Prok\_topoisomase.

CC Pfam: PF01131; Topoisom\_bac; 1.

CC Pfam: PF01131; Toprim; 1.

CC PRINTS: PRO0417; PRTPTSMRASEI.

CC SMART: SMO0437; TOP1ac; 1.

CC SMART: SMO0436; TOP1bc; 1.

CC SMART: SMO0435; TOP1c; 1.

CC SMART: SMO0434; TOP1d; 1.

CC SMART: SMO0433; TOP1e; 1.

DR TIGRPRMS; TIGR01051; topa\_bact; 1.

DR PROSITE; PS00396; TOPOISOMERASE\_I\_PROK; 1.

KW Isomerase; topoisomerase; DNA-binding; Zinc-finger; Metal-binding;

KW Complete proteome.

FT ZN FING 559 580 C4-TYPE.

FT ACT SITE 288 288 DNA CLEAVAGE (BY SIMILARITY).

SO SEQUENCE 633 AA; 72694 MW; F7262A044060CFE9 CRC64;

Query Match 15.1%; Score 61.5; DB 1; Length 633;

Best Local Similarity 25.4%; Pred. No. 12;

Matches 16; Conservative 14; Mismatches 32; Indels 1; Gaps 1;

QY 18 IKRVNVIHGRIDIPFASKFHLQKNTIGRRISIKLTKGSKASRRVDGIYAQNPA 77

DB 96 IARVNTLGRKNRI-VFSEITPRVIREAVKPNREIDMKVAQAQARRLDRIIVGYSISP 154

QY 78 SWK 80

DB 155 LWR 157

RESULT 14

RS18 DROME STANDARD; PRT; 152 AA.

AC P41094; Q9V911;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 40S ribosomal protein S18.

GN RPS18 OR CG8900.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

FN [1]

RP SEQUENCE FROM N.A. (LONG ISORPM).

RC STRAIN=Canton-S;

RX MEDLINE=94215909; PubMed=8163194;

RA Garwood J., Lepesant J.-A.,

RT "The Drosophila melanogaster homolog of ribosomal protein S18.";

RL Gene 141:231-235 (1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D., Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballwatz R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brockstein P., Brodtier P., Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dauter L., Davert L.B., Davies P., DePablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson R.C., Dougan L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M., Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP REVISIONS, AND ALTERNATIVE SPLICING.  
 RC STRAIN-BERKELEY;  
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochman S.E., Smith C.D.,  
 RA Tuzi J.L., Bergman C.M., Berman B.P., Carlson J.W., Celisner S.E.,  
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,  
 RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,  
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutnak F.,  
 RA Whitefield E.O., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE  
 CC INITIATION OF TRANSLATION (BY SIMILARITY).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a  
 CC short form; are produced by alternative splicing.  
 CC -1- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL, L22959; AAA28870.1; -;  
 DR EMBL, AE003792; AAM68401.1; -;  
 DR EMBL, AE003792; AAM68402.1; -;  
 DR Flybase; FBgn0010411; RPS18.  
 DR InterPro; IPR001892; Ribosomal\_S13.  
 DR Pfam; PF00416; Ribosomal\_S13; 1.  
 DR Prodom; PD001363; Ribosomal\_S13; 1.  
 DR PROSITE; PS00646; RIBOSOMAL\_S13; 1.  
 DR Ribosomal protein; Alternative splicing.  
 KM VARSPLIC 58  
 FT ECTEEVDKVTIISNPLOKYPNNELNROKDIIDGKYWL  
 FT DILQGLLENA (IN SHORT ISOFORM).  
 FT TSS -> RQGGDHLEPSAVQAGLVPOAGGHRKQVLLAA  
 FT VARSPLIC 102 152  
 FT MISSING (IN SHORT ISOFORM).  
 SQ SEQUENCE 152 AA; 17611 MW; F52P9E865B880EF CRC64;  
 QY Query Match 15.0%; Score 61; DB 1; Length 152;  
 QY Best Local Similarity 25.6%; Pred. No. 2.8; Gaps 3;  
 QY Matches 21; Conservative 10; Mismatches 29; Indels 22; Gaps 3;  
 DB 12 PNRVFIKRVSVNYIHGR---IDIFASKNFHLQKNTIGTGR-----ISIKLTS 58  
 DB 6 PEKQHIKRLMTNINIDKRVGIAMTAIK-----GVRKYVNIIVKKADVDTKRA 56  
 QY 59 GKIASRRVDGIIAAYONPASMK 80  
 DB 57 GECTEEVDKVTIISNPLOK 78  
 RESULT 15  
 ID Y12K\_SMSV4 STANDARD; PRT; 110 AA.  
 AC P36289;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Possible 12 kDa nucleic acid-binding protein.  
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).  
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 CC Vestivirus  
 OK NCBI\_TaxID=36407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion virus: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -1- SIMILARITY: TO FELINE CALICIVIRUS 12 kDa PROTEIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL, M87482; AAA16221.1; -;  
 DR PIR; D48562; D48562.  
 SQ SEQUENCE 110 AA; 12566 MW; 1425SD593827418E CRC64;  
 QY Query Match 14.9%; Score 60.5; DB 1; Length 110;  
 QY Best Local Similarity 32.4%; Pred. No. 2.2;  
 QY Matches 11; Conservative 10; Mismatches 12; Indels 1; Gaps 1;  
 DB 17 FIKRVSVNYIHGRDIIDIFASKNFHLQKNTIGTGR 50  
 DB 10 FLNSVANAVEGKLDL-ASKGLQKSRALDYER 42

Search completed: April 20, 2003, 13:07:25  
 Job time : 17.5789 secs

GenCore version 5.1.4 p5 4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2003, 12:57:30 ; Search time 32.4211 Seconds

(without alignments)  
237.215 Million cell updates/sec

Title: US-09-142-524D-1

Perfect score: 406

Sequence: 1 MKYTVAFNPGFNRVPIKR.....IASRRVDGIITAYQNPASWK 80

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	99	24.4	514	2	S48730	Cry j II protein -
2	99	24.4	514	2	JC2498	second major aller
3	74	18.2	514	2	JC7100	polygalacturonase
4	71	17.5	374	2	JC2124	major allergen Cry
5	71	17.5	374	2	JC2123	major allergen Cry
6	71	17.5	374	2	JC2123	major allergen Cry
7	70	17.2	152	1	S30393	ribosomal protein
8	70	17.2	152	1	R3RT18	ribosomal protein
9	70	17.2	152	1	R3RT18	ribosomal protein
10	68	16.7	394	2	T49115	pectate lyase like
11	67	16.5	674	2	F83794	ABC transporter (A
12	66.5	16.4	196	2	S54580	pectate lyase like
13	66	16.3	350	2	T49116	pectate lyase like
14	66	16.3	350	2	T49116	pectate lyase like
15	66	16.3	350	2	T49116	pectate lyase like
16	65.5	16.1	391	2	F89828	conserved hypochet
17	65.5	16.1	391	2	F72552	conserved hypochet
18	65.5	16.1	391	2	F72552	conserved hypochet
19	65	16.0	390	2	H86253	hypothetical prote
20	65	16.0	390	2	H86253	hypothetical prote
21	64.5	15.9	498	2	UC3353	hypothetical prote
22	64	15.8	222	2	AH2465	hypothetical prote
23	64	15.8	222	2	AH2465	hypothetical prote
24	63.5	15.6	433	2	F82298	hypothetical prote
25	63.5	15.6	433	2	F82298	hypothetical prote
26	62.5	15.4	533	2	T34458	conserved hypochet
27	62.5	15.4	533	2	T34458	conserved hypochet
28	62.5	15.4	533	2	T34458	conserved hypochet
29	62	15.3	368	2	A90266	heterodisulfide re

## ALIGNMENTS

30	62	15.3	5627	2	C83339	hypothetical prote
31	61.5	15.1	369	2	E75620	hypothetical prote
32	61.5	15.1	633	2	S62737	DNA topoisomerase
33	61	15.0	173	2	B27743	pape fibrillar prot
34	60.5	14.9	110	2	D48562	hypothetical nucle
35	60.5	14.9	856	2	B81399	hypothetical prote
36	60	14.8	154	2	T27228	probable periplasm
37	60	14.8	219	2	C84647	ribosomal protein
38	60	14.8	275	2	T28738	probable synaptoch
39	60	14.8	295	2	S60711	hypothetical prote
40	60	14.8	410	2	H90304	band-6-protein - b
41	60	14.8	627	2	T00484	conserved hypochet
42	60	14.8	725	2	S60712	hypothetical prote
43	60	14.8	844	1	TVMSV	band-6-protein - h
44	59.5	14.7	254	2	D69545	transforming prote
45	59.5	14.7	263	2	A72349	ABC transporter, A
						conserved hypochet

## RESULT 1

S48730  
Cry j II protein - Japanese cedar  
C:Species: Cryptomeria japonica (Japanese cedar)  
C:Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 28-May-1999

C:Accession: S48730  
R:Nanda, M.; Torioge, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Kur  
FBS Lett. 353, 124-128, 1994

A:Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar p  
A:Reference number: S48730; MUID:95010777; PMID:7926035

A:Accession: S48730  
A:Status: preliminary

A:Molecule type: mRNA  
A:Residues: 1-514 <NM>

A:Cross-references: GB:D37765; NID:G577695; PIDN:BA07021.1; PID:dl007598; PID:G577696

Query Match 24.4%; Score 99; DB 2; Length 514;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49  
DB 236 IDIFASKNFHLQKNTIGTG 254

## RESULT 2

JC2498  
second major allergen Cry j II precursor - Japanese cedar

C:Species: Cryptomeria japonica (Japanese cedar)  
C:Date: 16-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 21-Jul-2000

C:Accession: JC2498; PC2346; A60147  
R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994  
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese  
A:Reference number: JC2498; MUID:94271186; PMID:8002972

A:Accession: JC2498  
A:Molecule type: mRNA

A:Residues: 1-514 <KM>  
A:Cross-references: DDBJ:D29772; NID:G506857; PIDN:BA06172.1; PID:G506858

A:Accession: PC2346  
A:Molecule type: Protein

A:Residues: 52-61 <KO2>  
R:Sakaguchi, M.; Inouye, S.; Tanai, M.; Ando, S.; Usui, M.; Matuhasi, T.  
Allergy 45, 309-312, 1990

A:Title: Identification of the second major allergen of Japanese cedar pollen.  
A:Reference number: A60147; MUID:90342988; PMID:2382797

A:Accession: A60147  
A:Molecule type: protein

A:Residues: 55-64 <SAK>  
C:Keywords: glycoprotein; pollen  
F:1-54/Domain: signal sequence #status predicted <SIG>  
F:55-460/Product: second major allergen Cry j #status predicted <WAT>

F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.4%; Score 99; DB 2; Length 514;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49  
DB 236 IDIFASKNFHLQKNTIGTG 254

# RESULT 3

JC7100  
polygalacturonase Cha o 2 - Japanese cypress  
C/Species: Chamaecyparis obtusa (Japanese cypress)  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-May-2000  
C/Accession: JC7100; PC7026  
R/Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.  
Biochem. Biophys. Res. Commun. 263, 166-171, 1999  
A/Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all  
A/Reference number: JC7100; MUID:99417540; PMID:10486272  
A/Accession: JC7100  
A/Molecule type: protein  
A/Residues: 1-514 <MOR>  
A/Accession: PC7026  
A/Molecule type: protein  
A/Residues: 51-62 <MOR>

Query Match 18.2%; Score 74; DB 2; Length 514;  
Best Local Similarity 73.7%; Pred. No. 0.87;  
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49  
DB 236 IDIFASKNFHLQKNTIGTG 254

# RESULT 4

JC2124  
major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar  
C/Species: Cryptomeria japonica (Japanese cedar)  
C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000  
C/Accession: JC2124  
R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.  
Biochem. Biophys. Res. Commun. 199, 619-625, 1994  
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese  
A/Reference number: JC2123; MUID:94183224; PMID:8135802  
A/Accession: JC2124  
A/Molecule type: mRNA  
A/Residues: 1-374 <SON>  
A/Cross-references: GB:D26545; NID:9493633; PIDN:BA05543.1; PID:9493634  
A/Experimental source: pollen  
A/Note: the authors described carbohydrate binding site for residue 279  
C/Superfamily: peccate lyase LATS9  
C/Keywords: glycoprotein; pollen  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-374/Product: major allergen Cry j I (clone pCCI-15) #status predicted <MAT>  
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.5%; Score 71; DB 2; Length 374;  
Best Local Similarity 55.6%; Pred. No. 1.4;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAFNQFGPNRRVPIKRVSNVITH 27  
DB 234 MKVTVAFNQFGPNQGRMPRARVGLVH 260

# RESULT 5

JC2123  
major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar  
C/Species: Cryptomeria japonica (Japanese cedar)  
C/Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 21-Jul-2000

C/Accession: JC2123; PC2065  
R/Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.

Biochem. Biophys. Res. Commun. 199, 619-625, 1994  
A/Title: Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese  
A/Reference number: JC2123; MUID:94183234; PMID:8135802  
A/Accession: JC2123

A/Molecule type: mRNA  
A/Residues: 1-374 <SON>  
A/Cross-references: GB:D26544; NID:9493631; PIDN:BA05542.1; PID:9493632  
A/Experimental source: pollen  
A/Accession: PC2065

A/Molecule type: protein  
A/Residues: 22-53,58-81,219-232,236-258,299-307,346-372 <SO2>  
A/Note: the authors described carbohydrate binding site for residue 279  
C/Superfamily: peccate lyase LATS9  
C/Keywords: glycoprotein; pollen

F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-374/Product: major allergen Cry j I (clone pCCI-2-2) #status predicted <MAT>  
F:158,191,293,354/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.5%; Score 71; DB 2; Length 374;  
Best Local Similarity 55.6%; Pred. No. 1.4;  
Matches 15; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKVTVAFNQFGPNRRVPIKRVSNVITH 27  
DB 234 MKVTVAFNQFGPNQGRMPRARVGLVH 260

# RESULT 6

JC7366  
Jun a 2 protein - mountain cedar  
C/Species: Juniperus ashei (mountain cedar)  
C/Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 17-Nov-2000  
C/Accession: JC7366; PC7093  
R/Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.  
Biochem. Biophys. Res. Commun. 275, 195-202, 2000  
A/Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all  
A/Reference number: JC7366  
A/Accession: JC7366  
A/Molecule type: mRNA  
A/Residues: 1-507 <YOK>  
A/Cross-references: GB:AJ404653  
A/Accession: PC7093  
A/Molecule type: protein  
A/Residues: 55-63 <YOK>  
A/Comment: This protein, a second major allergen of mountain cedar pollen, which is invo  
o the polygalacturonase family.  
C/Keywords: glycoprotein; pollen

Query Match 17.5%; Score 71; DB 2; Length 507;  
Best Local Similarity 73.7%; Pred. No. 1.9;  
Matches 14; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 31 IDIFASKNFHLQKNTIGTG 49  
DB 237 IDIFASKNFHLQKNTIGTG 255

# RESULT 7

S30393  
ribosomal protein S18, cytosolic - human  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C/Accession: S30393; S68935  
R/Chassin, D.; Bellet, D.; Koman, A.  
Nucleic Acids Res. 21, 745, 1993  
A/Title: The human homolog of ribosomal protein S18.  
A/Reference number: S30393; MUID:93181276; PMID:8441687  
A/Accession: S30393

A/Molecule type: mRNA  
A/Residues: 1-152 <CHA>  
A/Cross-references: EMBL:X69150; NID:938422; PIDN:CA856794.1; PID:96006558



R.Vladimirov, S.N.; Ivanov, A.V.; Karpova, G.G.; Musolyanov, A.K.; Egorov, T.A.; Thiede, E. J. Biochem. 239, 144-149, 1996  
 A>Title: Characterization of the human small-ribosomal-subunit proteins by N-terminal an  
 A:Reference number: S68911; MUID:96305378; PMID:8706659  
 A:Accession: S68935

A:Molecule type: protein  
 A:Residues: 5', 56-57, 'X', 59-64, 'XX', 67-69 <VLA>  
 C:Genetics:

A:Gene: GDB:RPS18  
 A:Cross-references: GDB:138447; OMIM:180473

A:Map position: 6p21.3-6p21.3  
 C:Superfamily: Escherichia coli ribosomal protein S13  
 C:Keywords: blocked amino end; protein biosynthesis; ribosome

Query Match 17.2%; Score 70; DB 1; Length 152;  
 Best Local Similarity 30.4%; Pred. No. 0.68;

Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

QY 12 PNRVFIKRVSNVTHGRIDIFASKNFHLQKNTGTGR-----ISLKLTSGKI 61  
 DB 6 PEKQHLRLVLTNTIDGRKIAFAITAIK-----GVGRYAHVVLKADIDLTGRAGEL 59

QY 62 ASRRVDGIIAAYONPASMK 80  
 DB 60 TEDEVERVITIMONPROYK 78

## RESULT 8

R3RT18  
 ribosomal protein S18, cytosolic [validated] - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 21-Jul-2000

C:Accession: JH0419  
 R:Chan, Y.L.; Paz, V.; Wool, I.G.  
 Biochem. Biophys. Res. Commun. 178, 1212-1218, 1991

A>Title: The primary structure of rat ribosomal protein S18.  
 A:Reference number: JH0419; MUID:91337062; PMID:1872840  
 A:Accession: JH0419

A:Molecule type: mRNA  
 A:Residues: 1-152 <CHA>

A:Cross-references: GB:X57529; NID:9433446; PIDN:CAA0750.1; PID:9433447  
 A>Note: the protein is designated as ribosomal protein S18 according to conigration anal  
 C:Superfamily: Escherichia coli ribosomal protein S13  
 C:Keyword: protein biosynthesis; ribosome

Query Match 17.2%; Score 70; DB 1; Length 152;  
 Best Local Similarity 30.4%; Pred. No. 0.68;

Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

QY 12 PNRVFIKRVSNVTHGRIDIFASKNFHLQKNTGTGR-----ISLKLTSGKI 61  
 DB 6 PEKQHLRLVLTNTIDGRKIAFAITAIK-----GVGRYAHVVLKADIDLTGRAGEL 59

QY 62 ASRRVDGIIAAYONPASMK 80  
 DB 60 TEDEVERVITIMONPROYK 78

## RESULT 9

176666  
 ribosomal protein S18 [similarity] - mouse  
 N:Alternate names: ribosomal protein S13 [misnomer]

C:Species: Mus musculus (house mouse)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 03-Nov-2000

C:Accession: I76666; I57006  
 R:MacMurray, A.J.; Shin, H.S.  
 Mamm. Genome 2, 87-95, 1992

A>Title: The murine MHC encodes a mammalian homolog of bacterial ribosomal protein S13.  
 A:Reference number: I57006; MUID:92182530; PMID:1543907  
 A:Accession: I76666

A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-152 <RES>

A:Cross-references: GB:M76763; NID:9198579; PIDN:AAA16795.1; PID:9198580

A:Accession: I57006  
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-103, 'H', 105-152 <RE2>  
 A:Cross-references: GB:M76762; NID:9198577; PIDN:AAA16796.1; PID:9198578

C:Genetics:  
 A:Gene: Ke-3  
 A:Introns: 1/3, 34/3, 97/3, 128/2

C:Superfamily: Escherichia coli ribosomal protein S13

Query Match 17.2%; Score 70; DB 2; Length 152;  
 Best Local Similarity 30.4%; Pred. No. 0.68;

Matches 24; Conservative 7; Mismatches 32; Indels 16; Gaps 2;

QY 12 PNRVFIKRVSNVTHGRIDIFASKNFHLQKNTGTGR-----ISLKLTSGKI 61  
 DB 6 PEKQHLRLVLTNTIDGRKIAFAITAIK-----GVGRYAHVVLKADIDLTGRAGEL 59

QY 62 ASRRVDGIIAAYONPASMK 80  
 DB 60 TEDEVERVITIMONPROYK 78

## RESULT 10

T49115  
 pectate lyase like protein - Arabidopsis thaliana  
 N:Alternate names: protein AT4g22080

C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Sep-2000  
 C:Accession: T49115

R:Bevan, M.; Medler, H.; Mambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M  
 submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25016  
 A:Accession: T49115

A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-394 <BRV>

A:Cross-references: EMBL:AL022140; GSPDB:GNO0062; ATSP:AT4g22080  
 A:Experimental source: cultivar Columbia; BAC clone F1N20

C:Genetics:  
 A:Gene: ATSP:AT4g22080  
 A:Map position: 4

A:Introns: 240/3; 275/2; 321/3  
 C:Superfamily: pectate lyase LATS9

Query Match 16.7%; Score 68; DB 2; Length 394;  
 Best Local Similarity 28.6%; Pred. No. 3.3;

Matches 24; Conservative 7; Mismatches 27; Indels 26; Gaps 3;

QY 1 MKTVAFNPGPNRRVTKRVSVTH--GRRID-----IFASKNFHLQKN 44  
 DB 256 MKTVAFNPGPNRRVTKRVSVTH--GRRID-----IFASKNFHLQKN 44

QY 45 TIGTRRISLKLTSKGKIASRRVDG 68  
 DB 316 -----KSNKSKYTKREYKNG 329

## RESULT 11

F83794  
 ABC transporter (ATP-binding protein) BH1158 [imported] - Bacillus halodurans (strain C-7)

C:Species: Bacillus halodurans  
 C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 17-May-2002  
 C:Accession: F83794

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
 Nucleic Acids Res. 28, 4317-4331, 2000  
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132  
 A:Accession: F83794  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-674 <STO>

Query Match	16.5%	Score 67	DB 2	Length 674
Best Local Similarity	29.7%	Pred: No. 7.9		
Matches 19; Conservative	15	Mismatches 20	Indels 10	Gaps 3

```

QY      1  MKYTVAFNGEGRNRYFIKRVSNVTHGRIDIPASNPFHQNTIGIGARISLKLXISG 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     158  LMVVASFPQYG-QRFYLLQKAAKRRIIGRLRIDLFN----HLSRLPV----RFFDNMPAGK 207

```

QY	61	IASR	64
Db	208	VWSR	211

## RESULT 12

554580 probable membrane protein YMR118c - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: hypothetical protein YM9718.17c

CIDate: 08-Jul-1995 #sequence\_revision 19-Oct-1995 #text\_change 19-Apr-2002  
C:\Accession: S54580  
R:\Hunt, S.: Bowman, S.  
submitted to the EMBL Data Library, May 1995

A;Reference number: S54510  
A;Accession: S54580  
A;Molecule type: DNA  
A;Residues: 1-196 <HUN>

A;Cross-references: EMBL:Z49702; NID:gb17859; PID:gb17878; GSPDB:GN00013; MIPS:YMR1180  
A;Experimental source: strain AB972

C;Genetics:  
A;Gene: MIPS:YMR118c  
A;Cross-references: SGD:S0004724

A;Map position: 13R  
C;Keywords: transmembrane protein  
F;98-114/Domain: transmembrane #s  
F;179-195/Domain: transmembrane #

Query Match	16.4%	Score 66.5;	DB 2;	Length 196;
Best Local Similarity	29.7%;	Pred. No. 2.3;		
Matches 27;	Conservative 16;	Mismatches 33;	Indels 15;	Gaps 7;

```

OY      1 MKTV--AFNOC-PNRVFIKVS-----NYIHGRIDIPASKFH---LQKNTIG 48
        || : : || : : || : : || : : || : : || : : || : : || : :
Db      1 MKALIQVTSFGVPRASVVERISTFILLANTISG-RNDLF-SKEFHNGRSKSDWMS 58

```

```

QY      49 GRRISLKTSGKIASRVDGIIAAYNPASW 79
          : | | : | : | : |
Db      59 SNKEELLVSR-KKRPISPHLTVEEEMSW 88

```

## RESULT 13

S75065  
sensory transduction histidine kinase sll1590 - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: protein sll1590  
C:Species: *Synechocystis* sp.

A/Variety: PCC 6803  
C/Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C/Accession: S75065  
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.;

O, K.; Okumura, S.; Shimpō, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Issued DNA Res. 3, 109-135, 1996

A;Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* 5.

A/Reference number: S74322; MUID: 97061201; PMID: 8905231  
A/Accession: S75065  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-350 <KAN>

A1Cross-references: EMBL:ID09010; GB:AB001339; NID:gl552956; PIN:BAA11927.1; PID:d101866  
A1Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C1Genetics:  
A1Start codon: GTG

Query Match	16.3%	Score 66;	DB 2;	Length 350;
Best Local Similarity	39.5%	Pred. No. 5.1;		
Matches 17; Conservative	7;	Mismatches 15;	Indels 4;	Gaps 2.

```
QY      14 RRVFIKRVSVITIH---GRIDI-FASKNFILQKNTGTGTGARI 52
      :|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      246 QRFENNLITNALINHSPRGRKVEISLSSKNHFPQVQIVDEGRGI 288
```

## RESULT 14

pectate lyase like protein - Arabidopsis thaliana  
T49116  
N; Alternate names: protein AT4G22090  
O: Glucanase; Arabidopsis thaliana (mouse-ear cress)

CIspecies: *Manulobus thalassius* var. *crater*  
 C.Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
 C.Accession: T49116  
 R.Beyan, M.; Medler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.;  
 Submitted to the Detroit Commerce Database May 2000

A/Reference number: Z25016  
A/Accession: T49116  
A/Status: preliminary

A: Molecular type: DNA  
A: Residues: 1-394 <BEV>  
A: Cross-references: EMBL:AL022140; GSPDB:GNO0062; ATSP:AT4G22090  
A: Experimental source: cultivar Columbia; BAC clone FIN20

C;Genetics:  
A;Gene: AT4g22090  
A;Map position: 4  
A;Introns: 240/3; 275/

```
C:\superfamily: peccate lyase traits
Query Match      16.3%; Score 66; DB 2; Length 394;
Best Local Similarity 28.6%; Pred. No. 5.8;
Number of matches 27; Index 26; Gaps 3
```

[illegible]

D6 256 MKVIVAFNHFGFGLVEKMPKVRKGIATV  
QY 45 TIGTRRISLKLTSCKIASRRVDG 68  
| | : | |  
TGVGVETLVNEDVTC 220

## RESULT 15

P89982 conserved hypothetical protein SA0553 [imported] - Staphylococcus aureus (strain N315)  
C1:Species: Staphylococcus aureus  
C1:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C;Accession: F89828  
R;Kurroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
ma, A.; Mizumati-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.  
A;Reference number: A89758; PMID:11418146  
A;Accession: F89828

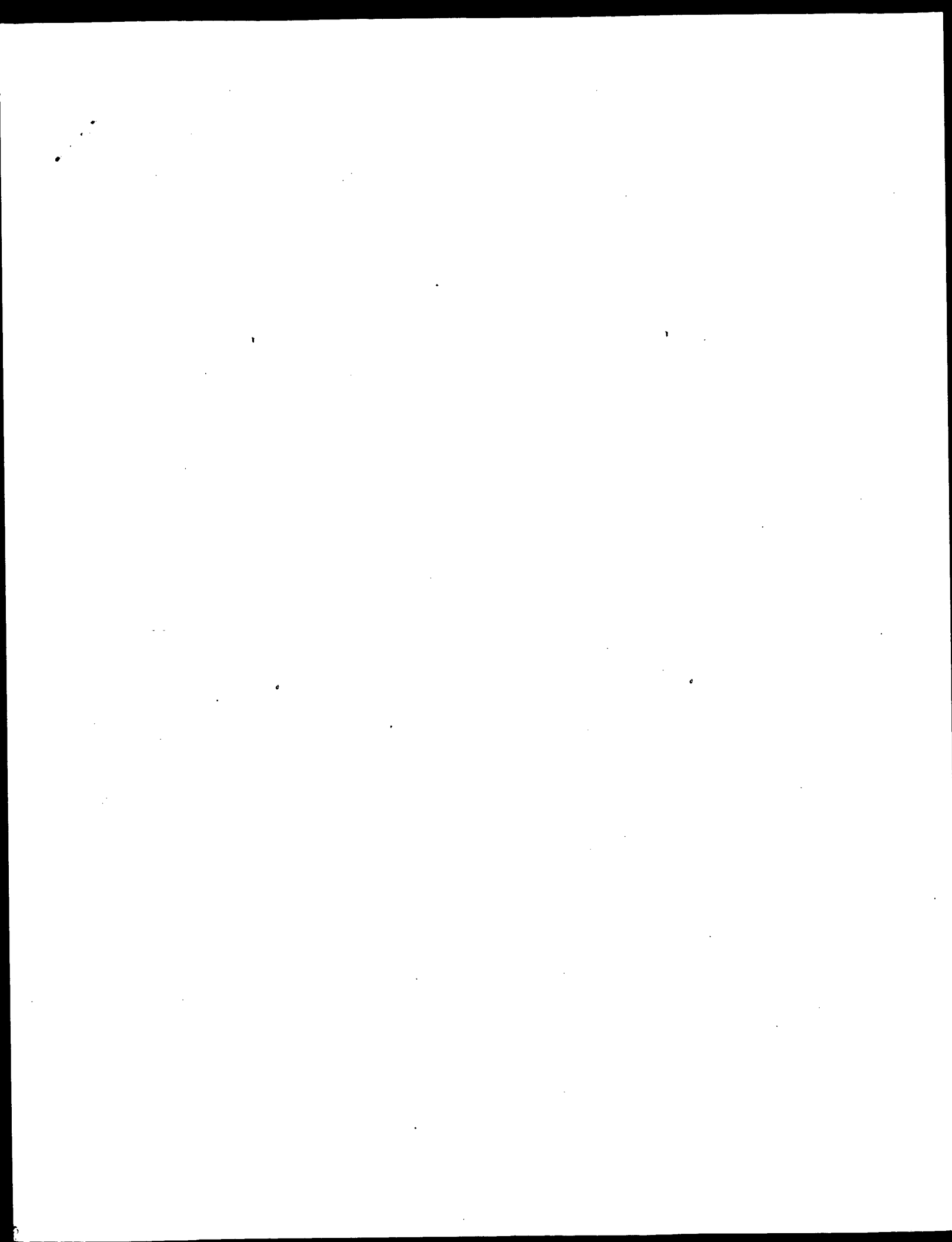
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-464 <KUR>  
A;Cross-references: GB:BA000018; PDB:1G13700488; PDB:1BA41785.1; GSPDB:GN0014

A;Experimental  
C;Genetics:  
A;Gene: SA0553

Query Match	16.3%	Score 66;	DB 2;	length 464;
Best Local Similarly	28.6%	Pred. No. 6.9;		
Matches	26;	Conservative	9;	Mismatches 28; Gaps 5

Qy 11 GNNRRVFIRVS-NVLI-HGRRIDIFASNNPHLOKNTIGTGRRI----- 52  
Db 302 GNSFVEVKRTSGNVFIENGSK---FAKGGKYLGGVAGVGFQGMDDDLANDNKTFGEA 358  
Qy 53 ----SLKLTSGKIASRRVDGIIAAYQNPASW 79  
Db 359 LSHNGMTLLAGSVAA---FGVAALVSNPVGW 386

Search completed: April 20, 2003, 13:15:34  
Job time : 35.4211 secs



> O <  
O | 0 < Intelligent

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-97.res made by jdelaval on Wed 23 Apr 103 17:47:55-PDT.

Query sequence being compared: US-09-142-524D-97 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-97 (1-15) with:  
File : 09142524.dep

```

100-
N -
U - 50-
M -
B -
E -
R -
O - 10-
S -
Q - 5-
U -
E -
C -
S -
SCORE 0 2 3 5 7 8 10 12 13 15
STDEV

```

## PARAMETERS

Similarity matrix Unitary K-tuple 2  
Mismatch penalty 1 Joining penalty 20  
Gap penalty 1.00 Window size 15  
Gap size penalty 0.05  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean 15 Median 16 Standard Deviation 0.00  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00  
Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

3 100% similar sequences to the query sequence were found:

Sequence Name	Description	Init. Opt.	Length	Score	Score	Sig.	Frame
---------------	-------------	------------	--------	-------	-------	------	-------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 15 0.00 0  
2. US-09-142-524D-2 Sequence 2, Application U 105 15 15 15 0.00 0  
3. US-09-142-524D-1 Sequence 1, Application U 80 15 15 15 0.00 0

1. US-09-142-524D-97 (1-15)  
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Daitiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKTVAFNQGPNRRVFIRKSVNVIHGRRIDIFASKNPHLQKNTIGTGRMKNRIMWQFAKLTGFTLMGR  
10 20 30 40 50 60 70

RPLMIISGNMNIKLKMPYIAGKYTFEDGRRAEVSYVHWGAKFIRRVGIIIAAYONPASWK  
80 90 100 110 120 130 X  
X  
VGIIIAAYONPASWK  
X (97)

2. US-09-142-524D-97 (1-15)

US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Daitiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
PRIOR FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKTVAFNQGPNRRVFIRKSVNVIHGRRIDIFASKNPHLQKNTIGTGRMKNRIMWQFAKLTGFTLMGR  
10 20 30 40 50 60 70

X 10 X  
VDGIIAAYONPASWK  
|||  
80 90 100 X  
RLKMPYIAGYKTFDGRVVDGIIAAYONPASWK

3. US-09-142-524D-97 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Some, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO: 1  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

X  
VDGIIAA  
|||  
10 20 30 40 50 60 70  
MKYTVANOGCPNRRVFIKRVSNVITHGRIDIFASKNPHLOKNTIGTGRRISLKLTSKRIASRRVDGIIAA

10 X  
YONPASWK  
|||  
YONPASWK  
80

> O <  
O | 0 IntellGenetics  
> O <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-100.res made by jdelaval on Wed 23 Apr 103 17:48:12-PDT.

Query sequence being compared: US-09-142-524D-100 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-100 (1-15) with:  
File : 09142524.pep

100-  
N -  
U 50-  
M -  
B -  
E -  
R -  
O -  
F 10-  
S -  
E 5-  
Q -  
U -  
N -  
N -  
C -  
E -  
S -  
SCORE 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15  
STDDEV -1

## PARAMETERS

Similarity matrix Unitary 1 K-tuple 2  
Mismatch penalty 1 Joining penalty 20  
Gap penalty 1.00 Window size 15  
Gap size penalty 0.05  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean Median Standard Deviation  
11 4 6.93  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

2 100% similar sequences to the query sequence were found:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
---------------	-------------	--------	------------------	------------

The list of other best scores is:

Sequence Name	Description	Length	Init. Opt. Score	Sig. Frame
1. US-09-142-524D-3 Sequence 3, Application U		134	15	15 0.58 0
2. US-09-142-524D-2 Sequence 2, Application U		105	15	15 0.58 0

\*\*\* 1 standard deviation below mean \*\*\*  
3. US-09-142-524D-1 Sequence 1, Application U 80 3 5 -1.15 0

1. US-09-142-524D-100 (1-15)  
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.58  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKYTVAFNPGPNRRVPIKRVSVIIHGRRIDIFASKNFHLOKNTTGTGRKNNRWLQFAKLTGFTLMGR  
10 20 30 40 50 60 70  
RPLMIIFSGNNIKLMPMYIAGYKTFDGRBRAEVSVHVNGAKFTRVDGIIAAYONPASMK  
80 90 100 110 120 130

2. US-09-142-524D-100 (1-15)  
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica





Initial Score = 13    Optimized Score = 13    Significance = 0.69



> O <  
O | O Intelligenetics  
> O <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-120.res made by jdelaval on Wed 23 Apr 103 17:48:52-PDT.

Query sequence being compared: US-09-142-524D-120 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-120 (1-15) with:  
File : 09142524.pep

```

100-
N -
U - 50-
M -
B -
E -
R -
O -
F - 10-
S -
E - 5-
Q -
U -
E -
N -
C -
S -
SCORE 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
STDEV

```

## PARAMETERS

Similarity matrix Unitary K-tuple  
Mismatch penalty 1 Joining penalty 2  
Gap penalty 1.00 Window size 20  
Gap size penalty 0.05 15  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean Median Standard Deviation  
14 15 0.00  
Times: CPU Total Elapsed  
00:00:00.00 00:00:00.00  
Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt.	Sign.	Frame
---------------	-------------	--------	-------	------	-------	-------

1. US-09-142-524D-1 Sequence 1, Application U	80	14	14	0.00	0
2. US-09-142-524D-2 Sequence 2, Application U	105	14	14	0.00	0
3. US-09-142-524D-3 Sequence 3, Application U	134	14	14	0.00	0

1. US-09-142-524D-120 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 0.00  
Residue Identity = 93% Matches = 14 Mismatches = 1  
Gaps = 0 Conservative Substitutions = 0

MKTVAPNGFRRRVFIKRVSNVILHGRIDIFASKNFHLOKNTIGTRRISLKLTSGLIASRRVDGIIA  
10 20 30 40 50 60 70  
X 10 X  
GIDIFASKNFHLOKNTIGTRRISLKLTSGLIASRRVDGIIA

YONPASKWK  
80

2. US-09-142-524D-120 (1-15)  
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 0.00  
Residue Identity = 93% Matches = 14 Mismatches = 1  
Gaps = 0 Conservative Substitutions = 0

X 10 X

GIDIFASKNFHLQKN  
 10 20 30 40 X 50 60 70  
 MKVTVAENQFGPNRRVFIRKVSNTIIGRRIDIFASKNFHLQKNITGTRRWKNNRIWLOFALITGFTLMGR  
 80 90 100  
 RLKMWNTIAGYKTFDGRVRVDGIIAAYONPASWK

3. US-09-142-524D-120 (1-15)  
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Some, Toshio  
 APPLICANT: Kume, Akinori  
 APPLICANT: Dairiki, Kazuo  
 APPLICANT: Iwama, Akiko  
 APPLICANT: Kino, Kohsuke  
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
 FILE REFERENCE: SPO-103  
 CURRENT APPLICATION NUMBER: US/09/142,524D  
 CURRENT FILING DATE: 1998-09-09  
 PRIOR APPLICATION NUMBER: PCT/JP97/00740  
 PRIOR FILING DATE: 1997-03-10  
 NUMBER OF SEQ ID NOS: 174  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 3  
 LENGTH: 134  
 TYPE: PRT  
 ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 0.00  
 Residue Identity = 93% Matches = 14 Mismatches = 1  
 Gaps = 0 Conservative Substitutions = 0

X 10 X  
 GIDIFASKNFHLQKN  
 10 20 30 40 X 50 60 70  
 MKVTVAENQFGPNRRVFIRKVSNTIIGRRIDIFASKNFHLQKNITGTRRWKNNRIWLOFALITGFTLMGR  
 80 90 100 110 120 130  
 RPLMTIFSGNNINIKLKMFWIAGYKTFDGRVRVDGIIAAYONPASWK

> 0 <  
O | 0 Intelligence  
> 0 <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524D-121.res made by jdelaval on Wed 23 Apr 103 17:49:11 PDT.

Query sequence being compared: US-09-142-524D-121 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff:

Results of the initial comparison of US-09-142-524D-121 (1-15) with:  
File: 09142524.pep

```

100-
N -
U - 50-
M -
B -
E -
R -
O -
F - 10-
S -
E - 5-
Q -
U -
N -
C -
E -
S -
SCORE 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
STDEV

```

## PARAMETERS

Similarity matrix Unitary K-tuple 2  
Mismatch penalty 1 Joining penalty 20  
Gap penalty 1.00 Window size 15  
Gap size penalty 0.05  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean 15 Median 16 Standard Deviation 0.00  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

3 100% similar sequences to the query sequence were found:

Sequence Name	Description	Init. Opt.	Length Score	Score	Sig. Frame
---------------	-------------	------------	--------------	-------	------------

1. US-09-142-524D-3 Sequence 3, Application U	134	15	15	0.00	0
2. US-09-142-524D-2 Sequence 2, Application U	105	15	15	0.00	0
3. US-09-142-524D-1 Sequence 1, Application U	80	15	15	0.00	0

1. US-09-142-524D-121 (1-15)  
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiyo  
APPLICANT: Kuno, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 134

ORGANISM: Cryptosporidia japonica  
TYPE: PRT

Initial Score = 15 Optimized Score = 15 Significance = 0.00  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

```

MKVTVAFNQFGPNRRVFIKRVSNVITHGRIDIPASKNPHLOKNTIGTGRMRWNNRITLQPAKLTGFTLMGR
10 20 30 40 50 60 70
RPLWTFSGNNMILKMPMYIAGYKTFDGRRAEVSYYVNGAKFIRRVGGIIIAAYNPASWK
80 90 100 110 120 130

```

2. US-09-142-524D-121 (1-15)  
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiyo  
APPLICANT: Kuno, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptosporidia japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

```

X 10 X
ASKNPHLOKNTIGT
|||||

```

MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRMNQNRIMLQFAKLTGFTLMGR  
 10 20 30 40 50 60 70  
 RLKXPMYIAGYKTFDGRVDCIIAAYQNPASWK  
 80 90 100

3. US-09-142-524D-121 (1-15)  
 US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Some, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiho

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.00

Residue Identity = 100% Matches = 15 Mismatches = 0

Gaps = 0 Conservative Substitutions = 0

10 20 30 40 50 60 70  
 X X  
 ASKNEHLQKNTIGT  
 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRMNQNRIMLQFAKLTGFTLMGR  
 10 20 30 40 50 60 70  
 YQNPASWK

> O <  
O | | O < Intelligenetics  
> O <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-131.res made by jdelaval on Wed 23 Apr 103 17:49:28-PDT.

Query sequence being compared: US-09-142-524D-131 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-131 (1-15) with:  
File : 09142524.pep

```

100-
N -
U - 50-
M -
B -
E -
R -
O -
F - 10-
S -
S -
Q -
U -
E -
N -
C -
E -
S -
SCORE 0 1 2 3 5 6 8 9 11 12 14
STDEV 0 0 0 0 0 0 0 0 0 0 0 0

```

## PARAMETERS

Similarity matrix Unitary K-tuple 2  
Mismatch penalty 1 Joining penalty 20  
Gap penalty 1.00 Window size 15  
Gap size penalty 0.05  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean 6 Median 3 Standard Deviation 6.93  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt. Frame
1. US-09-142-524D-3	Sequence 3, Application US/09142524D	134	14	1.15 0
2. US-09-142-524D-1	Sequence 1, Application US/09142524D	80	2	4 -0.58 0
3. US-09-142-524D-2	Sequence 2, Application US/09142524D	105	2	4 -0.58 0

1. US-09-142-524D-131 (1-15)  
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:  
APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 14 Optimized Score = 14 Significance = 1.15  
Residue Identity = 93% Matches = 14 Mismatches = 1  
Gaps = 0 Conservative Substitutions = 0

MSKVTAFNPGNRRVFIKRVSNVTHRRIDIPASKNFHLQKNKTGTGRKNNRIMQLQFALKTGPTLMGR  
10 20 30 40 50 60 70

X  
SRAEVSYYHVGAKF  
RPLWTFSGNNMTKLKMPMYIAGYKTFDGRRAEVSYYHVGAKFIRRVDDIIAAYONPASWK  
80 90 100 X 110 X 120 130

2. US-09-142-524D-131 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:  
APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 4 Significance = -0.58  
Residue Identity = 26% Matches = 4 Mismatches = 11  
Gaps = 0 Conservative Substitutions = 0

X 10 X  
SRAEVSYVHNGAKF  
10 X 20 30 40 50 60 70  
MKVTVAFNGFGPNRRVETIKRVSNVLIHGRRIDIFASKNPHLOKNTIGTGRWNNRIWLOFAKLTGFTLMGR

YONPASWK  
80

3. US-09-142-524D-131 (1-15)  
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 4 Significance = -0.58  
Residue Identity = 26% Matches = 4 Mismatches = 11  
Gaps = 0 Conservative Substitutions = 0

X 10 X  
SRAEVSYVHNGAKF  
10 X 20 30 40 50 60 70  
MKVTVAFNGFGPNRRVETIKRVSNVLIHGRRIDIFASKNPHLOKNTIGTGRWNNRIWLOFAKLTGFTLMGR  
RLKMPWTIAGYKTFDGRGRVDGIIAAYONPASWK  
80 90 100



> O <  
O | 0 Intelligenetics  
> O <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-152.res made by jdelaval on Wed 23 Apr 103 17:49:45-PDT.

Query sequence being compared: US-09-142-524D-152 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-152 (1-15) with:  
File : 09142524.pep

```

100-
N
U 50-
M
B
E
R
O 10-
S
E 5-
O
U
E
N
C
B
S
0
SCORE 0 1 3 4 5 7 8 9 11 12
STDEV 0 1 3 4 5 7 8 9 11 12

```

## PARAMETERS

Similarity matrix Unitary 1 K-tuple 2  
Mismatch penalty 1.00 Joining penalty 20  
Gap size penalty 0.05 Window size 15  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean 6 Median 4 Standard Deviation 5.20  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Opt.	Sig.	Frame
---------------	-------------	--------	-------	------	------	-------

1. US-09-142-524D-1 Sequence 1, Application U	**** 1 standard deviation above mean ****	80	12	12	1.15	0
2. US-09-142-524D-2 Sequence 2, Application U	**** 0 standard deviation from mean ****	105	3	4	-0.58	0
3. US-09-142-524D-3 Sequence 3, Application U		134	3	4	-0.58	0

1. US-09-142-524D-152 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Some, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiyo  
APPLICANT: Kuno, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 12 Optimized Score = 12 Significance = 1.15  
Residue Identity = 80% Matches = 12 Mismatches = 3  
Gaps = 0 Conservative Substitutions = 0

```

MKVTVAFNQGFENRERFVKRVSNVLIHGRIDIFASKNFHLQKNTIGTGRRISLKTSGKIASRNVGIIAA
10 20 30 40 50 60 70
YQNPASWK 80

```

2. US-09-142-524D-152 (1-15)  
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Some, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiyo  
APPLICANT: Kuno, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 3 Optimized Score = 4 Significance = -0.58  
Residue Identity = 27% Matches = 5 Mismatches = 10  
Gaps = 3 Conservative Substitutions = 0

X 10 X  
LSDISL---KLTSGKIAS  
| | | |  
10 20 30 40 50 60 70

MKVTVAFNPGFGRNRVFIKRVSNVHIGRRIDIFASKNFHLOKNTIGTRGRMKNNRIWLOFAKLTGFTLMGR  
10 20 30 40 50 60 70  
RLKMPMYIAGYKTFDGRRVGIIAAYQNPASWK  
80 90 100

3. US-09-142-524D-152 (1-15)  
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kobsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Cryptosporidia japonica

Initial Score = 3 Optimized Score = 4 Significance = -0.58  
Residue Identity = 27% Matches = 5 Mismatches = 10  
Gaps = 3 Conservative Substitutions = 0

X 10 X  
LSDISL---KLTSGKIAS  
| | | |  
10 20 30 40 50 60 70  
MKTVAFNPGFGRNRVFIKRVSNVHIGRRIDIFASKNFHLOKNTIGTRGRMKNNRIWLOFAKLTGFTLMGR  
80 90 100 110 120 130  
RPLMIIFSGNMNIKLMKMPYIAGYKTFDGRRAEVSIVHVGAKFIRVDGIIAAYQNPASWK

> O <  
O | 10 Intelligenetics  
> O <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-28.res made by jdelaval on Wed 23 Apr 103 17:45:42-PDT.

Query sequence being compared: US-09-142-524D-28 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-28 (1-15) with:  
File : 09142524.pep

```

100-
N -
U - 50-
M -
B -
E -
R -
O - 10-
F -
S -
Q - 5-
U -
E -
N -
C -
S -
S 0-----
SCORE 0 2 3 5 7 8 10 12 13 15
STDEV 0 1 1 1 1 1 1 1 1 1

```

## PARAMETERS

Similarity matrix Unitary 1 K-tuple 2  
Mismatch penalty 1 Joining penalty 20  
Gap penalty 1.00 Window size 15  
Gap size penalty 0.05  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean Median Standard Deviation  
6 3- 7.51  
Times: CPU  
00:00:00.00 Total Elapsed  
00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length Score	Score	Sig.	Frame
---------------	-------------	------------	--------------	-------	------	-------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 1.20 0

The list of other best scores is:

Sequence Name	Description	Length Score	Score	Sig.	Frame
2. US-09-142-524D-1 Sequence 1, Application U		80	2	3	-0.53 0
3. US-09-142-524D-2 Sequence 2, Application U		105	2	3	-0.53 0

1. US-09-142-524D-28 (1-15)

US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunoth-  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 1.20  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKTVAVNGSPNRRVRVIRKVSNTVHGRRIDIFASKNFHLQKNTGTGRMRKNNRIMLQFAKLTGFTLMGR  
10 20 30 40 50 60 70

X 10 X  
RPLWIFSGNNMILK  
RPLWIFSGNNMILKPKMTIAGKTFDGRRAEVSVHVNGAKFIRVDGIIAAYONPASMK  
X 80 X 90 100 110 120 130

2. US-09-142-524D-28 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 80  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 3 Significance = -0.53

*Young I*  
*Venus Seq 1D*  
*# 1, 2, 3*

jic Diseases

Residue Identity = 23% Matches = 4 Mismatches = 11  
 Gaps = 2 Conservative Substitutions = 0

X  
 RPLWIFSGN--NNIKL  
 X  
 10  
 MKTVAFNQGPPRRRVFIKRVSNVITIGRRIDIFASKNFHLQKNTIGTGRRISLKTSGKIASRRVDGIIA  
 X 10 20 30 40 50 60 70  
 YONPASWK  
 80

3. US-09-142-524D-28 (1-15)  
 US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
 APPLICANT: Kume, Akio  
 APPLICANT: Dairiki, Kazuo  
 APPLICANT: Iwama, Akiko  
 APPLICANT: Kino, Kohsuke  
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
 FILE REFERENCE: SPQ-103

CURRENT APPLICATION NUMBER: US/09/142, 524D  
 CURRENT FILING DATE: 1998-09-09  
 PRIOR APPLICATION NUMBER: PCT/JP97/00740  
 PRIOR FILING DATE: 1997-03-10  
 NUMBER OF SEQ ID NOS: 174  
 SOFTWARE: Patent version 3.1  
 SEQ ID NO 2  
 LENGTH: 105

TYPE: PRT  
 ORGANISM: Cryptomeria japonica

Initial Score = 2 Optimized Score = 3 Significance = -0.53  
 Residue Identity = 23% Matches = 4 Mismatches = 11  
 Gaps = 2 Conservative Substitutions = 0

X  
 RPLWIFSGN--NNIKL  
 X  
 10  
 MKTVAFNQGPPRRRVFIKRVSNVITIGRRIDIFASKNFHLQKNTIGTGRRISLKTSGKIASRRVDGIIA  
 X 10 20 30 40 50 60 70  
 RLKMPYIAGYKTFDGRRRVDGIIAAYQNPASWK  
 80 90 100

> 0  
0/10 Intelligenetics  
> 0 <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-29.res made by jdelaval on Wed 23 Apr 103 17:46:13-PDT.

Query sequence being compared: US-09-142-524D-29 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-29 (1-15) with:  
File : 09142524.pep

100-  
N -  
U -  
M -  
B -  
E -  
R -  
O -  
F -  
S -  
E -  
C -  
N -  
U -  
E -  
N -  
N -  
C -  
E -  
S -  
SCORE 0 2 3 5 7 8 10 12 13 15  
STDEV 0 1 1 1 1 1 1 1 1 1

## PARAMETERS

Similarity matrix Unitary 2  
Mismatch penalty 1 K-tuple 20  
Gap penalty 1.00 Joining penalty 20  
Gap size penalty 0.05 Window size 15  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean Median Standard Deviation  
7 3 6.66  
Times: CPU  
00:00:00.00 Total Elapsed  
00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% similar sequence to the query sequence was found:

Sequence Name	Description	Init. Opt.	Length	Score	Sig. Frame
---------------	-------------	------------	--------	-------	------------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 1.20 0

The list of other best scores is:

Sequence Name	Description	Length	Score	Sig. Frame
2. US-09-142-524D-2 Sequence 2, Application U		105	6	7 -0.15 0
3. US-09-142-524D-1 Sequence 1, Application U		80	2	3 -0.75 0

1. US-09-142-524D-29 (1-15)

US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Some, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 3  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 1.20  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKVTVAENQFGPNRRVFIKRVSNVNIHGRRIDIPASKNPHLOKNTIGTGRWKNRIMLOPAKLTGFTLMGR  
10 20 30 40 50 60 70

X 10 X  
IFSGNNNITKIKPMY  
|||||  
RPLWTFSGNNNITKIKPMYIAGYKTFDGRRAEVSYYVANGAKFIRVDGIIAAYONPASWK  
80 90 X 100 110 120 130

2. US-09-142-524D-29 (1-15)

US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Some, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 2  
LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 6 Optimized Score = 7 Significance = -0.15

Residue Identity = 46% Matches = 7 Mismatches = 8  
 Gaps = 0 Conservative Substitutions = 0

X  
 IFSGNMNI

MKVTVAFNQFGPNRRVFIKRVSNVITIHGRIDIPASKNPHLQKNITIGTGRMRWNNRIWLOFAKLTGFTLMGR  
 10 20 30 40 50 60 70

10 X

KLKMPMY

|||||

RLKMPMYIAGYKTFDGRVDGIITAAQONPASKWK  
 80 90 100

### 3. US-09-142-524D-29 (1-15) US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akimori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740-

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 2 Optimized Score = 3 Significance = -0.75  
 Residue Identity = 20% Matches = 3 Mismatches = 12  
 Gaps = 0 Conservative Substitutions = 0

X 10 X  
 IFSGNMNITKLKMPMY

MKVTVAFNQFGPNRRVFIKRVSNVITIHGRIDIPASKNPHLQKNITIGTGRRIISLKLTSKIASRRVDGIITAA  
 10 20 X 30 40 50 60 70

YONPASKWK  
 80

LENGTH: 105  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score	=	11	Optimized Score	=	12	Significance	=	0.30
Residue Identity	=	80%	Matches	=	12	Mismatches	=	3
Gaps	=	0	Conservative Substitutions	=			=	0

MKVTVAFNQFGENRRVFYIKRVSNTYIHGRRLIDIPASNFHLOKNTIGTGRMKNRRITWLOFAKLTGTFTLMGR  
 10 20 30 40 50 60 70

10 X  
 KLKMPNYIAGYK  
 |||||  
 RLKMPNYIAGYKTFDGRVDCIIAAYQNPASWK  
 80 X 90 100

3. US-09-142-524D-30 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
 APPLICANT: Kume, Akinori  
 APPLICANT: Dairiki, Kazuo  
 APPLICANT: Iwama, Akiko  
 APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D  
CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: 08/110,687A

PRIOR APPLICATION NUMBER: PCT/JP97/007400  
 PRIOR FILING DATE: 1997-03-10  
 NUMBER OF SEQ ID NOS: 174

NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1

LENGTH: 80  
TYPE: PRT  
ORGANISM: Cryptomeria japonica

```
Initial Score = 2 Optimized Score = 4 Significance = -1.05
Residue Identity = 26% Matches = 4 Mismatches = 11
Gaps = 0 Conservative Substitutions = 0
```

X 10 X  
MNIKLMPIYAGYK

10 X 20 30 40 50 60 70

YONPASWK  
80



> 0 <  
0 | 0 Intelligenetics  
> 0 <

PaetDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-31.res made by jdelaval on Wed 23 Apr 103 17:46:55-PDT.

Query sequence being compared: US-09-142-524D-31 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-31 (1-15) with:  
File : 09142524.pep

100-  
N -  
U 50-  
M -  
B -  
E -  
R -  
O 10-  
S -  
O 5-  
U -  
E -  
N -  
C -  
S -  
SCORE 0 2 3 5 7 8 10 12 13 15  
STDEV

## PARAMETERS

Similarity matrix Unitary 1 K-tuple 2  
Mismatch penalty 1.00 Joining penalty 20  
Gap size penalty 0.05 Window size 15  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean Median Standard Deviation  
10 3 7.51  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00  
Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

2 100% similar sequences to the query sequence were found:

Sequence Name	Description	Int. Opt.	Length	Score	Sig. Frame
---------------	-------------	-----------	--------	-------	------------

1. US-09-142-524D-3 Sequence 3, Application U 134 15 15 0.67 0  
2. US-09-142-524D-2 Sequence 2, Application U 105 15 15 0.67 0  
The list of other best scores is:

Sequence Name	Description	Length	Score	Int. Opt.	Sig. Frame
---------------	-------------	--------	-------	-----------	------------

3. US-09-142-524D-1 Sequence 1, Application U 80 2 3 -1.07 0  
\*\*\*\* 1 standard deviation below mean \*\*\*\*

1. US-09-142-524D-31 (1-15)  
US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3

LENGTH: 134

TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.67  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKVTVAFNQGPNRRVFIKRVSNVLIHGRRIDIFASKNFHLQKNTIGTGRWKNRIMLOPAKLTGFTLMGR  
10 20 30 40 50 60 70

X 10 X  
KMPMYAGYKTFDGR  
RPLWTFSGNNMNIKLMKMPYTAGYKTFDGRRAEVSYYVHVGAKFIREVDGIIMAYONPASWK  
80 90 100 X 110 120 130

2. US-09-142-524D-31 (1-15)  
US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akimori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103  
CURRENT APPLICATION NUMBER: US/09/142,524D

CURRENT FILING DATE: 1998-09-09  
PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10  
NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2

LENGTH: 105

TYPE: PRT  
ORGANISM: Cryptomeria japonica

Initial Score = 15 Optimized Score = 15 Significance = 0.67  
Residue Identity = 100% Matches = 15 Mismatches = 0  
Gaps = 0 Conservative Substitutions = 0

MKVTVAFNQGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNNTIGTGRRWKNRIMLQFAKLTGFTLMGR  
10 20 30 40 50 60 70

X 10 X  
KMPMYIAGYKTFDGR  
|||||  
RLKMPMYIAGYKTFDGRVDGIITAAVQNPASWK  
X 80 90 100

3. US-09-142-524D-31 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 2 Optimized Score = 3 Significance = -1.07  
Residue Identity = 20% Matches = 4 Mismatches = 11  
Gaps = 5 Conservative Substitutions = 0

X 10 X  
KMPMYIAGYK-----TFDGR  
| |  
MKVTVAFNQGPNRRVFIKRVSNVLIHGRIDIFASKNFHLQKNNTIGTGRRIISLKTSGKIASRRVDGIITAA  
10 20 30 40 50 60 70  
YQNPASWK  
80

> 0 <  
O | O IntelliGenetics  
> 0 <

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-142-524d-36.res made by jdelaval on Wed 23 Apr 103 17:47:17-PDT.

Query sequence being compared: US-09-142-524D-36 (1-15)  
Number of sequences searched: 3  
Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-36 (1-15) with:  
File : 09142524.pep

```

100-
N -
U 50-
M -
B -
R -
O 10-
S -
E 5-
U -
E -
N -
E -
C -
E -
S 0-
SCORE 0 1 3 4 6 7 9 10 12 13
STDEV

```

## PARAMETERS

Similarity matrix Unitary 2  
Mismatch penalty 1 Joining penalty 20  
Gap penalty 1.00 Window size 15  
Gap size penalty 0.05  
Cutoff score 0  
Randomization group 0

## SEARCH STATISTICS

Scores: Mean 13 Median 14 Standard Deviation 0.00  
Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00

Number of residues: 319  
Number of sequences searched: 3  
Number of scores above cutoff: 3

The scores below are sorted by initial score.  
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Int. Opt. Score	Sig. Frame
---------------	-------------	--------	-------	-----------------	------------

1. US-09-142-524D-1	Sequence 1, Application U	80	13	13	0.00 0
2. US-09-142-524D-2	Sequence 2, Application U	105	13	13	0.00 0
3. US-09-142-524D-3	Sequence 3, Application U	134	13	13	0.00 0

1. US-09-142-524D-36 (1-15)  
US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiho

APPLICANT: Kuno, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 80

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00  
Residue Identity = 86% Matches = 13 Mismatches = 2  
Gaps = 0 Conservative Substitutions = 0

```

X 10 X
PCVFIRKRVSNVING
|||||
MKVTVAFNQFGNRRVFIKRVSNVILHGRIDIPASKNFHLQKNTIGRRRLSLKTSKIASRRVDGIIA
10 X 20 30 40 50 60 70

```

YONPASWK  
80

2. US-09-142-524D-36 (1-15)

US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Dairiki, Kazuo

APPLICANT: Iwama, Akiho

APPLICANT: Kuno, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

PRIOR FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 2

LENGTH: 105

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00  
Residue Identity = 86% Matches = 13 Mismatches = 2  
Gaps = 0 Conservative Substitutions = 0

X 10 X

PCVFIRKRVSNVLIHG  
 |||||  
 10 X 20 30 40 50 60 70  
 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRIIDIFASKNFHLQKNTIGTGRMKNNRIWLOFAKLIGFTLMGR  
 80 90 100  
 RLKMPMYIAGYKTFDGRRVDDGIIAAYQNPASWK

3. US-09-142-524D-36 (1-15)  
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D

GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kume, Akinori

APPLICANT: Daitiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kino, Kohsuke

TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases

FILE REFERENCE: SPO-103

CURRENT APPLICATION NUMBER: US/09/142,524D

PRIOR FILING DATE: 1998-09-09

PRIOR APPLICATION NUMBER: PCT/JP97/00740

NUMBER OF SEQ ID NOS: 174

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Cryptosporidia japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00  
 Residue Identity = 86% Matches = 13 Mismatches = 2  
 Gaps = 0 Conservative Substitutions = 0

X 10 X  
 PCVFIRKRVSNVLIHG  
 |||||  
 10 X 20 30 40 50 60 70  
 MKVTVAFNQFGPNRRVFIKRVSNVLIHGRIIDIFASKNFHLQKNTIGTGRMKNNRIWLOFAKLIGFTLMGR  
 80 90 100 110 120 130  
 RPLMIIFSGMNIKLKMPMYIAGYKTFDGRRAEVSYYHVGAKFIRRVDDGIIAAYQNPASWK

> O <  
 O | 0 IntelliGenetics  
 > O <  
 FastDB - Fast Pairwise Comparison of Sequences  
 Release 5.4

Results file us-09-142-524d-57.res made by jdelaval on Wed 23 Apr 103 17:47:36-PDT.

Query sequence being compared: US-09-142-524D-57 (1-15)  
 Number of sequences searched: 3  
 Number of scores above cutoff: 3

Results of the initial comparison of US-09-142-524D-57 (1-15) with:  
 File: 09142524.pep

```

100-
N -
U -
M -
B -
E -
R -
O -
F 10-
S -
E 5-
U -
Q -
N -
C -
E -
S 0-
SCORE 0 1 3 4 6 7 9 10 12 13
STDEV

```

## PARAMETERS

Similarity matrix Unitary K-tuple 2  
 Mismatch penalty 1 Joining penalty 20  
 Gap penalty 1.00 Window size 15  
 Gap size penalty 0.05  
 Cutoff score 0  
 Randomization group 0

## SEARCH STATISTICS

Scores: Mean 13 Median 14 Standard Deviation 0.00  
 Times: CPU 00:00:00.00 Total Elapsed 00:00:00.00  
 Number of residues: 319  
 Number of sequences searched: 3  
 Number of scores above cutoff: 3

The scores below are sorted by initial score.  
 Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Int. Opt. Score	Sig. Frame
---------------	-------------	--------	-------	-----------------	------------

1. US-09-142-524D-1	Sequence 1, Application U	80	13	13	0.00 0
2. US-09-142-524D-2	Sequence 2, Application U	105	13	13	0.00 0
3. US-09-142-524D-3	Sequence 3, Application U	134	13	13	0.00 0

1. US-09-142-524D-57 (1-15)  
 US-09-142-524D-1 Sequence 1, Application US/09142524D

Sequence 1, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
 APPLICANT: Kume, Akinori  
 APPLICANT: Daitiki, Kazuo  
 APPLICANT: Iwama, Akiko  
 APPLICANT: Kino, Kohsuke  
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
 FILE REFERENCE: SPO-103  
 CURRENT APPLICATION NUMBER: US/09/142,524D  
 CURRENT FILING DATE: 1998-09-09  
 PRIOR APPLICATION NUMBER: PCT/JP97/00740  
 PRIOR FILING DATE: 1997-03-10  
 NUMBER OF SEQ ID NOS: 174  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 1  
 LENGTH: 80  
 TYPE: PRT  
 ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00  
 Residue Identity = 100% Matches = 13 Mismatches = 0  
 Gaps = 0 Conservative Substitutions = 0

```

X 10 X
KSMKTVAFNFGPN
|||||
MKXTVAFNFGPNRRVFIKRVSVIIHGRIDIPASKNPHLQKNTIGTGRISLKLTSKIASRRVDGII
X 10 X 20 30 40 50 60 70

```

AAAYONPASWK  
 80

2. US-09-142-524D-57 (1-15)

US-09-142-524D-2 Sequence 2, Application US/09142524D

Sequence 2, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
 APPLICANT: Kume, Akinori  
 APPLICANT: Daitiki, Kazuo  
 APPLICANT: Iwama, Akiko  
 APPLICANT: Kino, Kohsuke  
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
 FILE REFERENCE: SPO-103  
 CURRENT APPLICATION NUMBER: US/09/142,524D  
 CURRENT FILING DATE: 1998-09-09  
 PRIOR APPLICATION NUMBER: PCT/JP97/00740  
 PRIOR FILING DATE: 1997-03-10  
 NUMBER OF SEQ ID NOS: 174  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 2  
 LENGTH: 105  
 TYPE: PRT  
 ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00  
 Residue Identity = 100% Matches = 13 Mismatches = 0  
 Gaps = 0 Conservative Substitutions = 0

X 10 X

KSMKVTVAFNQFGPN  
 |||||  
 MKVTVAFNQFGPNRVFIRKVSNIHGRIDIFASKNFHLQKNTIGTGRKWKNNRIWLOPAKLTGFTLM  
 X 10 X 20 30 40 50 60 70  
 80 90 100  
 GRRLKMPMYIAGYKTFDGRVVGIIAAYONPASMK

3. US-09-142-524D-57 (1-15)  
 US-09-142-524D-3 Sequence 3, Application US/09142524D

Sequence 3, Application US/09142524D  
 GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
 APPLICANT: Kume, Akimori  
 APPLICANT: Dairiki, Kazuo  
 APPLICANT: Iwama, Akiko  
 APPLICANT: Kino, Kohsuke  
 TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Diseases  
 FILE REFERENCE: SPO-103  
 CURRENT APPLICATION NUMBER: US/09/142, 524D  
 CURRENT FILING DATE: 1998-09-09  
 PRIOR APPLICATION NUMBER: PCT/JP97/00740  
 PRIOR FILING DATE: 1997-03-10  
 NUMBER OF SEQ ID NOS: 174  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Cryptomeria japonica

Initial Score = 13 Optimized Score = 13 Significance = 0.00  
 Residue Identity = 100% Matches = 13 Mismatches = 0  
 Gaps = 0 Conservative Substitutions = 0

X 10 X  
 KSMKVTVAFNQFGPN  
 |||||  
 MKVTVAFNQFGPNRVFIRKVSNIHGRIDIFASKNFHLQKNTIGTGRKWKNNRIWLOPAKLTGFTLM  
 X 10 X 20 30 40 50 60 70  
 80 90 100 110 120 130  
 GRRLKMPMYIAGYKTFDGRVVGIIAAYONPASMK

=> fil reg

FILE 'REGISTRY' ENTERED AT 18:25:31 ON 23 APR 2003  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file  
provided by InfoChem.

STRUCTURE FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9  
DICTIONARY FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

Please note that search-term pricing does apply when  
conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP  
PROPERTIES for more information. See STNote 27, Searching Properties  
in the CAS Registry File, for complete details:  
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

=> d sta que l16

L7 56 SEA FILE=REGISTRY ABB=ON PLU=ON (RPLWIIIFSGNMNKL)|(IFSGNMNKLKM  
PMY)|(MNIKLKMPMYIAGYK)|(KMPMYIAGYKTFDGR)|(PCVFKRVSNNVIIHG)|(KSMK  
VTVAFNQFGPN)/SQSP  
L15 24 SEA FILE=REGISTRY ABB=ON PLU=ON (VDGIIAAYQNPAWK)|(NNRIWLQFAK  
LTGFT)|(LQFAKLTGFTLMGKG)|(GIDIFASKNFHLQKN)|(ASKNFHLQKNTIGTG)|(S  
RAEVSYPVHVNGAKF)|(LSDISLKLTSKGIAS)/SQSP  
L16 6 SEA FILE=REGISTRY ABB=ON PLU=ON L7 AND L15

=> d l16 sqide can tot

L16 ANSWER 1 OF 6 REGISTRY COPYRIGHT 2003 ACS  
RN 502526-68-3 REGISTRY  
CN 24: PN: JP2003079389 SEQID: 1 unclaimed protein (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 105

PATENT ANNOTATIONS (PNTE):

Sequence |Patent  
Source |Reference

=====+=====

Not Given	JP2003079389
	unclaimed
	SEQID 1

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR  
=====

51 RWKNNRIWLQ FAKLTGFTLM GRRLKMPMYI AGYKTFDGRR VDGIIAAYQN  
=====

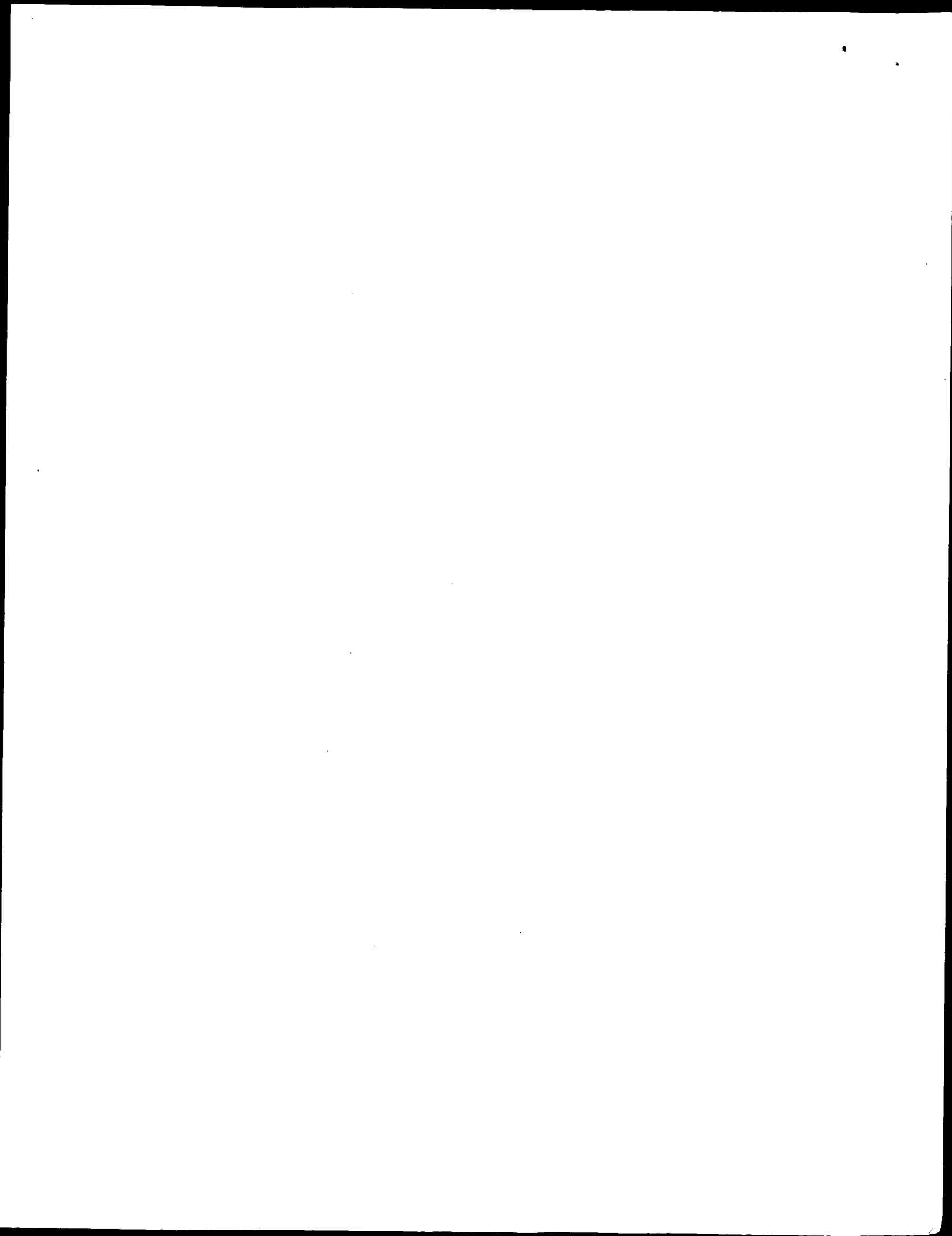
101 PASWK  
=====

HITS AT: 35-49, 54-68, 75-89, 91-105

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified  
CI MAN  
SR CA

Jan Delaval  
Reference Librarian  
Biology & Chemical Library  
Call 1-800-703-3003-4493  
[jan.delaval@uspis.gov](mailto:jan.delaval@uspis.gov)





LC STN Files: CA, CAPLUS  
1 REFERENCES IN FILE CA (1962 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:253826

L16 ANSWER 2 OF 6 REGISTRY COPYRIGHT 2003 ACS  
RN 502526-45-6 REGISTRY  
CN 1: PN: JP2003079389 SEQID: 3 unclaimed protein (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 134

PATENT ANNOTATIONS (PNTE):

Sequence |Patent  
Source |Reference

=====+=====

Not Given	JP2003079389
	unclaimed
	SEQID 3

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR  
=====

51RWKNNRIWLQ FAKLTGFTLM GRRPLWIIFS GNMNIKLKMP MYIAGYKTFD  
=====

101 GRRAEVSYVH VNGAKFIRRV DGIIAAYQNP ASWK  
=====

HITS AT: 35-49, 54-68, 83-102, 120-134

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS  
1 REFERENCES IN FILE CA (1962 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:253826

L16 ANSWER 3 OF 6 REGISTRY COPYRIGHT 2003 ACS  
RN 501062-84-6 REGISTRY  
CN Allergen (Cryptomeria japonica synthetic allergen multi-T cell epitope B)  
(9CI) (CA INDEX NAME)

OTHER NAMES:

CN 3: PN: JP2003079387 SEQID: 3 claimed protein  
FS PROTEIN SEQUENCE  
SQL 134

PATENT ANNOTATIONS (PNTE):

Sequence |Patent  
Source |Reference

=====+=====

Not Given	JP2003079387
	claimed
	SEQID 3

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR  
=====

51 RWKNNRIWLQ FAKLTGFTLM GRRPLWIIFS GNMNIKLKMP MYIAGYKTFD  
=====

101 GRRAEVSYVH VNGAKFIRRV DGIIAAYQNP ASWK  
=====



HITS AT: 35-49, 54-68, 83-102, 120-134

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS

1 REFERENCES IN FILE CA (1962 TO DATE)

1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:220362

L16 ANSWER 4 OF 6 REGISTRY COPYRIGHT 2003 ACS

RN 501062-83-5 REGISTRY

CN Allergen (Cryptomeria japonica synthetic allergen multi-T cell epitope A)  
(9CI) (CA INDEX NAME)

OTHER NAMES:

CN 1: PN: JP2003079387 SEQID: 1 claimed protein

FS PROTEIN SEQUENCE

SQL 105

PATENT ANNOTATIONS (PNTE):

Sequence |Patent

Source |Reference

=====+=====

Not Given|JP2003079387

|claimed

|SEQID 1

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR

51 RWKNNRIWLQ FAKLTGFTLM GRRLKMPMYI AGYKTFDGRV VDGIIAAYQN

101 PASWK

HITS AT: 35-49, 54-68, 75-89, 91-105

\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS

1 REFERENCES IN FILE CA (1962 TO DATE)

1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 138:220362

L16 ANSWER 5 OF 6 REGISTRY COPYRIGHT 2003 ACS

RN 196718-37-3 REGISTRY

CN Multiepitope allergen (synthetic Cryptomeria japonica allergen Cry j 1/Cry  
j 2-derived 134-amino acid isoform) (9CI) (CA INDEX NAME)

FS PROTEIN SEQUENCE

SQL 134

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR

51 RWKNNRIWLQ FAKLTGFTLM GRRPLWIIFS GNMNIKLKMP MYIAGYKTFD

101 GRRAEVSIVH VNGAKFIRRV DGIIAAYQNP ASWK

HITS AT: 35-49, 54-68, 83-102, 120-134



## \*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified  
CI MAN  
SR CA  
LC STN Files: CA, CAPLUS  
1 REFERENCES IN FILE CA (1962 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

L16 ANSWER 6 OF 6 REGISTRY COPYRIGHT 2003 ACS  
RN 196718-36-2 REGISTRY  
CN Multiepitope allergen (synthetic Cryptomeria japonica allergen Cry j 1/Cry  
j 2-derived 105-amino acid isoform) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 105

SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR  
=====   
51 RWKNNRIWLQ FAKLTGFTLM GRRLKMPMYI AGYKTFDGRR VDGIIAAYQN  
=====   
101 PASWK  
=====

HITS AT: 35-49, 54-68, 75-89, 91-105

## \*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\*

MF Unspecified  
CI MAN  
SR CA  
LC STN Files: CA, CAPLUS  
1 REFERENCES IN FILE CA (1962 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

=&gt; d his l16-

(FILE 'REGISTRY' ENTERED AT 18:15:33 ON 23 APR 2003)  
L16 6 S L7 AND L15  
SAV L7 DIBRINO1/A  
SAV L15 DIBRINO2/A  
SAV L16 DIBRINO3/A

FILE 'HCAOLD' ENTERED AT 18:22:48 ON 23 APR 2003  
L17 0 S L16

FILE 'HCAPLUS' ENTERED AT 18:22:56 ON 23 APR 2003  
L18 3 S L16  
SEL RN  
DEL SEL  
SEL RN 3

FILE 'REGISTRY' ENTERED AT 18:23:36 ON 23 APR 2003  
L19 5 S E1-E5  
L20 2 S L19 AND L16  
L21 3 S L19 NOT L20  
L22 1 S L21 AND L7,L15  
L23 2 S L21 NOT L22

FILE 'USPATFULL, USPAT2' ENTERED AT 18:25:15 ON 23 APR 2003  
L24 0 S L16



FILE 'REGISTRY' ENTERED AT 18:25:31 ON 23 APR 2003

=&gt; fil hcaplus

FILE 'HCAPLUS' ENTERED AT 18:25:56 ON 23 APR 2003

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 23 Apr 2003 VOL 138 ISS 17

FILE LAST UPDATED: 22 Apr 2003 (20030422/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

=&gt; d 118 all tot

L18 ANSWER 1 OF 3 HCAPLUS COPYRIGHT 2003 ACS

AN 2003:210113 HCAPLUS

DN 138:253826

TI Purification of recombinant multi-T cell epitope from inclusion body

IN Miyahara, Michinori; Mori, Takeshi; Yokoyama, Minehiko; Kamijo, Masayuki; Yamazaki, Tetsuya; Yamashita, Akio; Osamu, Kosuke

PA Meiji Milk Products, Co., Ltd., Japan

SO Jpn. Kokai Tokkyo Koho, 16 pp.

CODEN: JKXXAF

DT Patent

LA Japanese

IC ICM C12P021-02

ICS C12N015-09

CC 16-1 (Fermentation and Bioindustrial Chemistry)

Section cross-reference(s): 1, 3, 15

FAN.CNT 1

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
PI	JP 2003079389	A2	20030318	JP 2002-189240	20020628
PRAI	JP 2001-196579	A	20010628		

AB The recombinant multi-T cell epitope polypeptide is purified from inclusion body of Escherichia coli by solubilization with guanidine/urea; and subjection to copper-chelated chromatog., cation exchange chromatog., and reverse-phase chromatog. Purifn. of recombinant multi-T cell epitope polypeptide A comprising 6 epitopes of Cry J1 and j2, the main allergen protein of Cryptomeria japonica, by construction of plasmid pQTF7.DELTA.cr, transformation of E. coli, etc. was shown.

ST sugi pollen allergen protein epitope chromatog purifn; Escherichia recombinant sugi pollen allergen protein epitope

IT Allergens

RL: BPN (Biosynthetic preparation); PUR (Purification or recovery); BIOL (Biological study); PREP (Preparation)

(Cry j I (Cryptomeria japonica, I); purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Allergens





RL: BPN (Biosynthetic preparation); PUR (Purification or recovery); BIOL (Biological study); PREP (Preparation)  
 (Cry j II; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Affinity chromatography  
 (copper-chelated; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Plasmids  
 (pQTF7.DELTA.cr; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Cryptomeria japonica  
 (pollen; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Cation exchange chromatography  
 DNA sequences

Epitopes

Escherichia coli

Fermentation

Inclusion bodies

Pollen

Protein sequences

Reversed phase chromatography

(purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT Immunization  
 (vaccination; purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT 57-13-6, Urea, biological studies 113-00-8, Guanidine  
 RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)

(purifn. of recombinant multi-T cell epitope from inclusion body of Escherichia coli)

IT 502526-46-7 502526-47-8 502526-48-9 502526-49-0 502526-50-3  
 502526-51-4 502526-52-5 502526-53-6 502526-54-7 502526-55-8  
 502526-56-9 502526-57-0 502526-58-1 502526-59-2 502526-60-5  
 502526-61-6 502526-62-7 502526-63-8 502526-64-9 502526-65-0  
 502526-66-1 502526-67-2 502526-69-4

RL: PRP (Properties)

(unclaimed nucleotide sequence; purifn. of recombinant multi-T cell epitope from inclusion body)

IT 502526-45-6 502526-68-3

RL: PRP (Properties)

(unclaimed protein sequence; purifn. of recombinant multi-T cell epitope from inclusion body)

L18 ANSWER 2 OF 3 HCAPLUS COPYRIGHT 2003 ACS

AN 2003:210110 HCAPLUS

DN 138:220362

TI Expression of recombinant Japanese cedar allergen multi-T cell epitope in Escherichia coli inclusion body

IN Miyahara, Michinori; Mori, Takeshi; Yokoyama, Minehiko; Kamijo, Masayuki; Yamazaki, Tetsuya; Yamashita, Akio; Kino, Kouuke

PA Meiji Milk Products, Co., Ltd., Japan

SO Jpn. Kokai Tokkyo Koho, 19 pp.

CODEN: JKXXAF

DT Patent

LA Japanese

IC ICM C12N015-09

ICS A61K035-74; A61K048-00; A61P037-00

CC 15-2 (Immunochemistry)

Section cross-reference(s): 3, 16

FAN.CNT 1

PATENT NO.

KIND DATE

APPLICATION NO.

DATE



PI	JP 2003079387	A2	20030318	X	JP 2002-189243	20020628
PRAI	JP 2001-196598	A	20010628			

AB Recombinant expression of multi-T cell epitopes comprising 6 or 7 epitopes of major allergens of Japanese cedar (*C. japonica*) pollen, Cry j 1 and Cry j2, in *E. coli* as inclusion bodies, is disclosed. A construct contg. trp operon promoter/operator region, Shine-Dalgarno sequence, multi-T cell epitope coding sequence, and bacteriophage .lambda. terminator to, is used.

ST sugi pollen allergen epitope recombinant expression *Escherichia* inclusion body; T cell epitope *Cryptomeria* pollen allergen *Escherichia* inclusion body; Japanese cedar allergen epitope *Escherichia* inclusion body expression

IT Allergens  
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)  
(Cry j I (*Cryptomeria japonica*, I); expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Allergens  
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)  
(Cry j II; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT *Cryptomeria japonica*  
Epitopes  
*Escherichia coli*  
Inclusion bodies  
Pollen  
Protein sequences  
T cell (lymphocyte)  
cDNA sequences  
(expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Genetic element  
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)  
(operator, trp operon; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Genetic element  
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)  
(ribosome-binding site; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

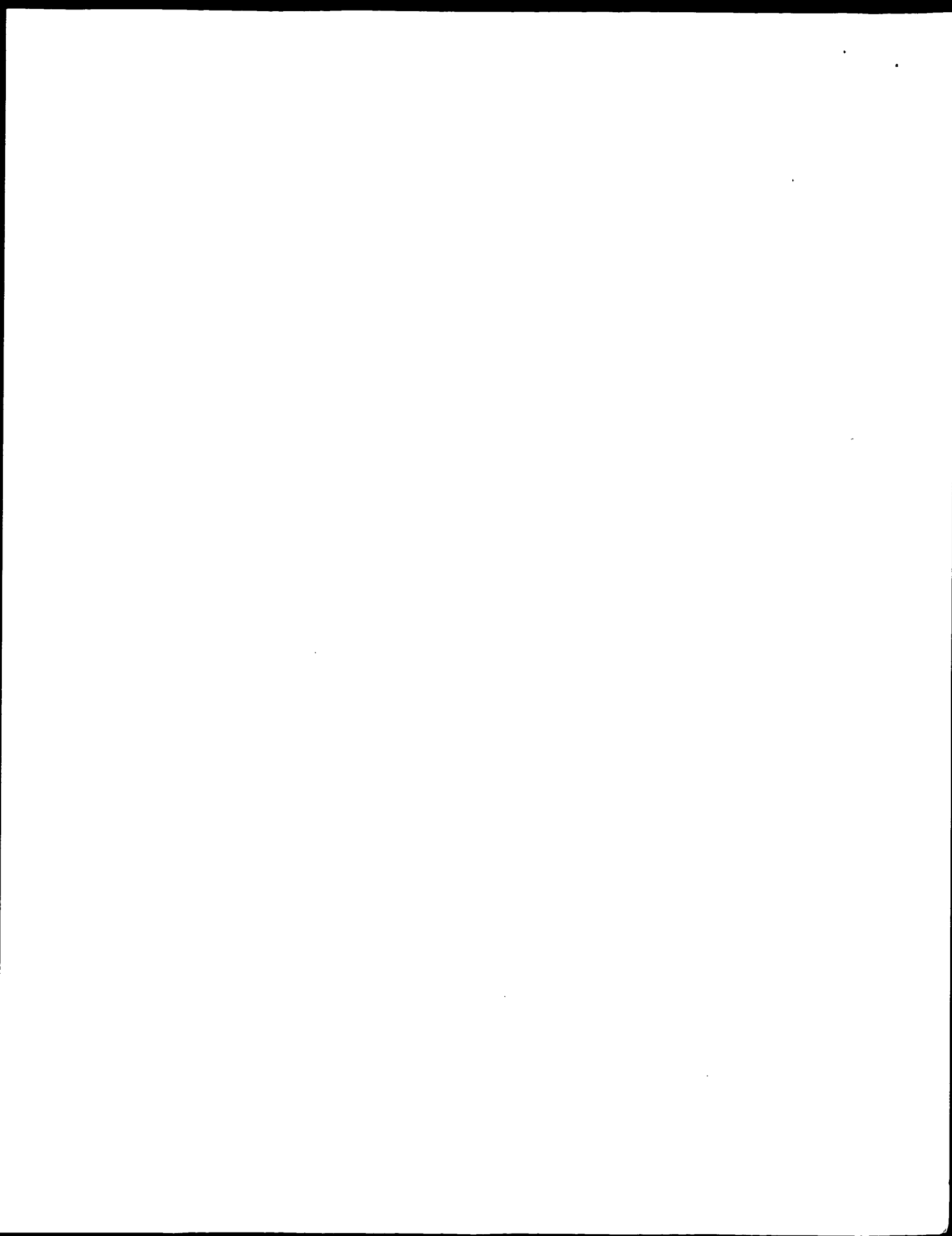
IT Coliphage .lambda.  
(terminator of; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Genetic element  
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)  
(terminator, bacteriophage .lambda.; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Promoter (genetic element)  
RL: BUU (Biological use, unclassified); BIOL (Biological study); USES (Uses)  
(trp operon; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT Operon  
(trp; expression of recombinant Japanese cedar allergen multi-T cell epitope in *Escherichia coli* inclusion body)

IT **501062-83-5P**  
RL: BPN (Biosynthetic preparation); PRP (Properties); BIOL (Biological study); PREP (Preparation)  
(amino acid sequence; expression of recombinant Japanese cedar allergen



multi-T cell epitope in Escherichia coli inclusion body)  
 IT 501062-84-6  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
 (Biological study)  
 (amino acid sequence; expression of recombinant Japanese cedar allergen  
 multi-T cell epitope in Escherichia coli inclusion body)  
 IT 501062-85-7  
 RL: BSU (Biological study, unclassified); BUU (Biological use,  
 unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)  
 (nucleotide sequence; expression of recombinant Japanese cedar allergen  
 multi-T cell epitope in Escherichia coli inclusion body)  
 IT 501062-86-8  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
 (Biological study)  
 (nucleotide sequence; expression of recombinant Japanese cedar allergen  
 multi-T cell epitope in Escherichia coli inclusion body)  
 IT 501070-61-7 501070-62-8 501070-63-9 501070-64-0 501070-65-1  
 501070-66-2 501070-67-3 501070-68-4 501070-69-5 501070-70-8  
 501070-71-9 501070-72-0 501070-73-1 501070-74-2 501070-75-3  
 501070-76-4 501070-77-5 501070-78-6 501070-79-7 501070-80-0  
 501070-81-1 501070-82-2  
 RL: PRP (Properties)  
 (unclaimed nucleotide sequence; expression of recombinant Japanese  
 cedar allergen multi-T cell epitope in Escherichia coli inclusion body)

L18 ANSWER 3 OF 3 HCAPLUS COPYRIGHT 2003 ACS

AN 1997:617989 HCAPLUS

DN 127:277192

TI Peptide-base immunotherapeutic agent for allergic diseases

IN Sone, Toshio; Kume, Akinori; Dairiki, Kazuo; Iwama, Akiko; Kino, Kohsuke

PA Meiji Milk Products Co., Ltd., Japan; Sone, Toshio; Kume, Akinori;

Dairiki, Kazuo; Iwama, Akiko; Kino, Kohsuke

SO PCT Int. Appl., 58 pp.

CODEN: PIXXD2

DT Patent

LA Japanese

IC ICM A61K039-35

ICS C07K014-725

CC 15-2 (Immunochemistry)

FAN.CNT 1

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9732600	A1	19970912	WO 1997-JP740	19970310
W: CA, CN, JP, KR, US				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
CA 2248937	AA	19970912	CA 1997-2248937	19970310
CN 1218412	A	19990602	CN 1997-194541	19970310
EP 923940	A1	19990623	EP 1997-906863	19970310
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				

PRAI JP 1996-80702 19960310

WO 1997-JP740 19970310

AB A peptide-base immunotherapeutic agent contg. an efficacious amt. of a monomol. multi-epitope peptide prepd. by bonding T-cell epitope regions derived from different allergen mols. with each other, e.g. Cry j 1/Cry i 2 or Cry j 1/Cha 0 1. It is efficacious for the prevention and therapy of a wide variety of allergic diseases.

ST allergen multiple epitope immunotherapy allergic disease

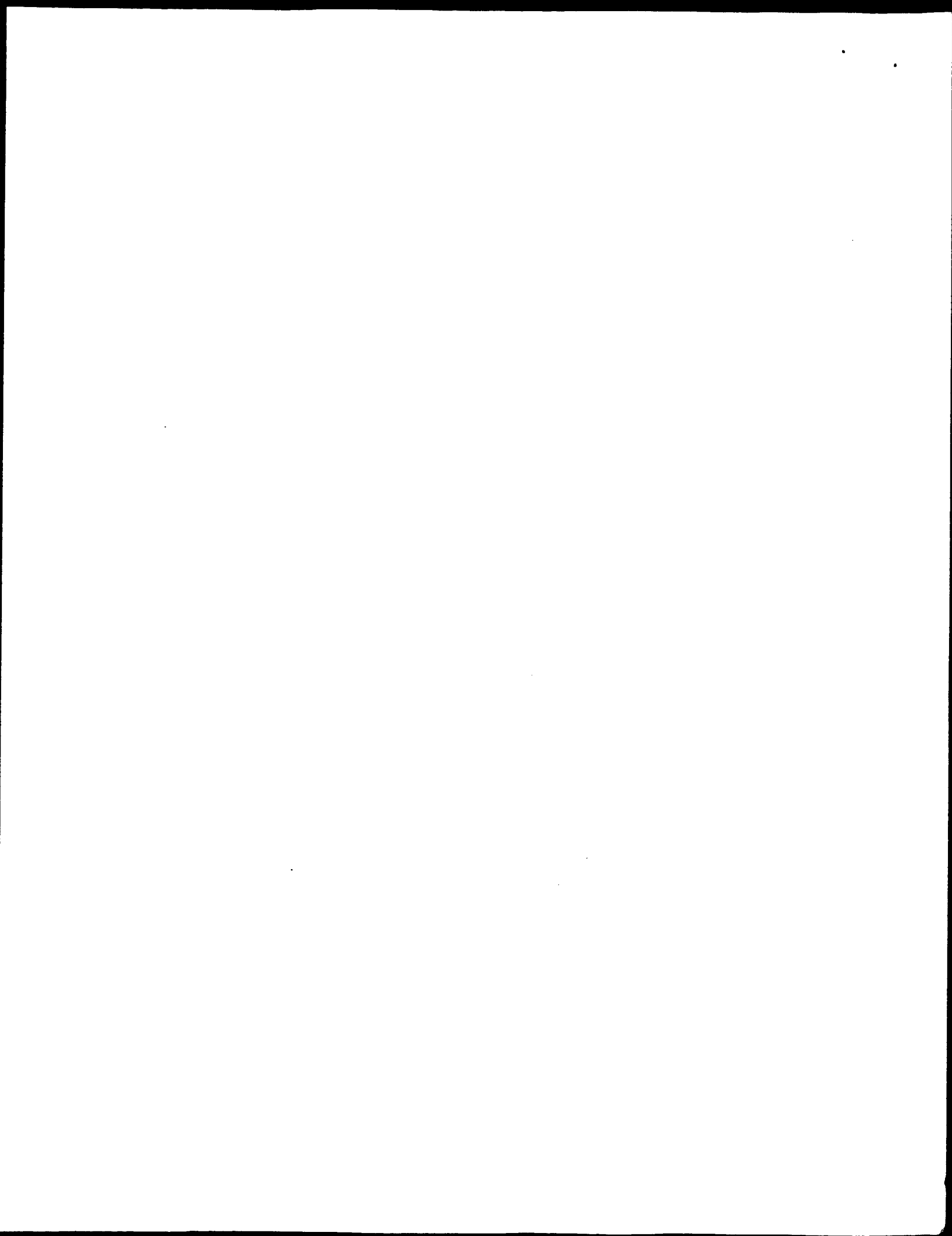
IT Allergens

RL: BSU (Biological study, unclassified); BIOL (Biological study)

(Cha 0 1 multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)



- IT Allergens  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(Cry j I (Cryptomeria japonica, I), multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Allergens  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(Cry j II, multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Immunoglobulins  
RL: BSU (Biological study, unclassified); BIOL (Biological study)  
(E; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens  
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)  
(HLA-DP, DPB1\*0501 and DPA1\*0101 and DPB1\*0201; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens  
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)  
(HLA-DQ, DQB1\*0602 and DQA1\*0102; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens  
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)  
(HLA-DR, DRB5\*0101 and DRB4\*0101; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Histocompatibility antigens  
RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)  
(HLA; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Epitopes  
RL: BSU (Biological study, unclassified); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
(T cell; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Antigen presentation  
(cells; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT T cell (lymphocyte)  
(epitopes; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Cryptomeria japonica  
Immunotherapy  
Lymphocyte  
Protein sequences  
(multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT Allergens  
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(multi-epitope; multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)
- IT 196718-35-1P 196718-36-2P 196718-37-3P  
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(amino acid sequence; multi-epitope peptides of Cry j 1 and/or Cry j 2





and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)

IT 196510-89-1P 196521-96-7P  
RL: BPN (Biosynthetic preparation); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); USES (Uses)  
(multi-epitope peptides of Cry j 1 and/or Cry j 2 and/or Cha 0 1 allergens are prepd. for immunotherapy of allergic diseases)

=> fil reg

FILE 'REGISTRY' ENTERED AT 18:26:09 ON 23 APR 2003  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
COPYRIGHT (C) 2003 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file provided by InfoChem.

STRUCTURE FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9  
DICTIONARY FILE UPDATES: 22 APR 2003 HIGHEST RN 503805-80-9

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 6, 2003

Please note that search-term pricing does apply when conducting SmartSELECT searches.

Crossover limits have been increased. See HELP CROSSOVER for details.

Experimental and calculated property data are now available. See HELP PROPERTIES for more information. See STNote 27, Searching Properties in the CAS Registry File, for complete details:  
<http://www.cas.org/ONLINE/STN/STNOTES/stnotes27.pdf>

=> d sqide can tot 121

L21 ANSWER 1 OF 3 REGISTRY COPYRIGHT 2003 ACS  
RN 196718-35-1 REGISTRY  
CN Multi-epitope allergen (synthetic Cryptomeria japonica allergen Cry j 1/Cry j 2-derived 80-amino acid isoform) (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 80  
  
SEQ 1 MKVTVAFNQF GPNRRVFIKR VSNVIIHGRR IDIFASKNFH LQKNTIGTGR  
51 RISLKLTSKG IASRRVDGII AAYQNPASWK  
MF Unspecified  
CI MAN  
SR CA  
LC STN Files: CA, CAPLUS  
1 REFERENCES IN FILE CA (1962 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

L21 ANSWER 2 OF 3 REGISTRY COPYRIGHT 2003 ACS  
RN 196521-96-7 REGISTRY  
CN L-Isoleucine, L-isoleucyl-L-phenylalanyl-L-seryl-L-lysyl-L-asparaginyl-L-leucyl-L-asparaginyl-L-isoleucyl-L-lysyl-L-leucyl-L-asparaginyl-L-methionyl-L-prolyl-L-leucyl-L-tyrosyl-L-isoleucyl-L-alanylglycyl-L-asparaginyl-L-lysyl-L-arginyl-L-arginyl-L-phenylalanyl-L-isoleucyl-L-lysyl-L-arginyl-L-valyl-L-seryl-L-asparaginyl-L-valyl- (9CI) (CA INDEX NAME)  
FS PROTEIN SEQUENCE  
SQL 31



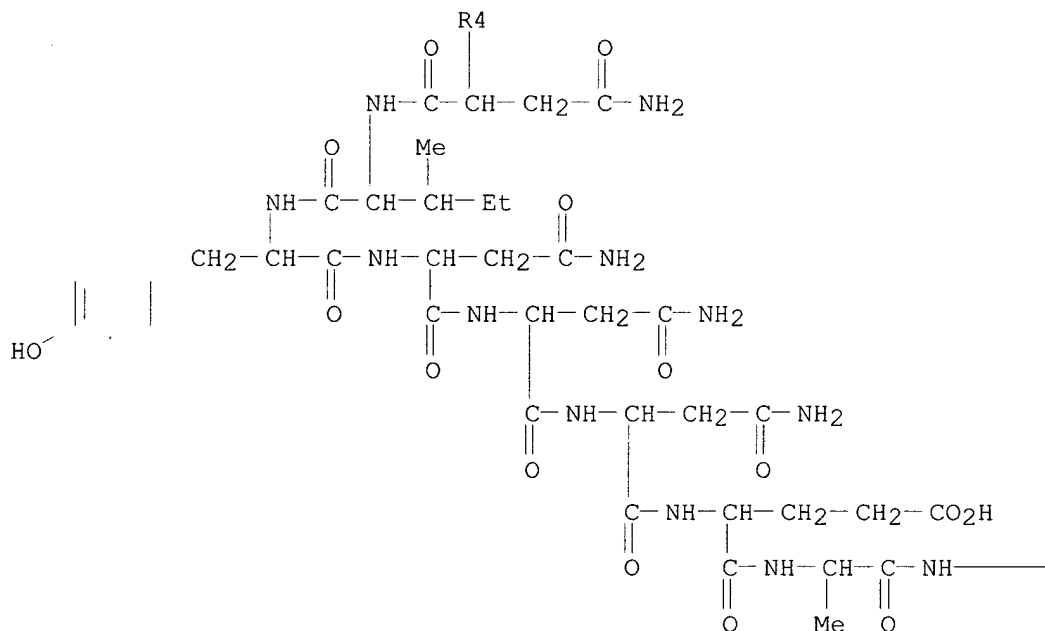
SEQ 1 IFSKNLNIKL NMPLYIAGNK RRFIKRVSNV I  
 MF C168 H283 N49 O40 S  
 CI MAN  
 SR CA  
 LC STN Files: CA, CAPLUS  
 1 REFERENCES IN FILE CA (1962 TO DATE)  
 1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

L21 ANSWER 3 OF 3 REGISTRY COPYRIGHT 2003 ACS  
 RN 196510-89-1 REGISTRY  
 CN L-Isoleucine, L-seryl-L-serylglycyl-L-lysyl-L-asparaginyl-L-.alpha.-  
 glutamylglycyl-L-threonyl-L-asparaginyl-L-isoleucyl-L-tyrosyl-L-  
 asparaginyl-L-asparaginyl-L-asparaginyl-L-.alpha.-glutamyl-L-alanyl-L-  
 phenylalanyl-L-lysyl-L-valyl-L-.alpha.-glutamyl-L-arginyl-L-arginyl-L-  
 phenylalanyl-L-isoleucyl-L-lysyl-L-arginyl-L-valyl-L-seryl-L-asparaginyl-L-  
 valyl- (9CI) (CA INDEX NAME)  
 FS PROTEIN SEQUENCE; STEREOSEARCH  
 SQL 31

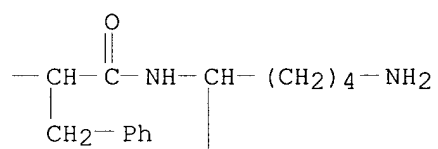
SEQ 1 SSGKNEGTNI YNNNEAFKVE RRFIKRVSNV I  
 MF C155 H251 N49 O49  
 SR CA  
 LC STN Files: CA, CAPLUS

PAGE 1-A

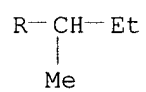




PAGE 1-B

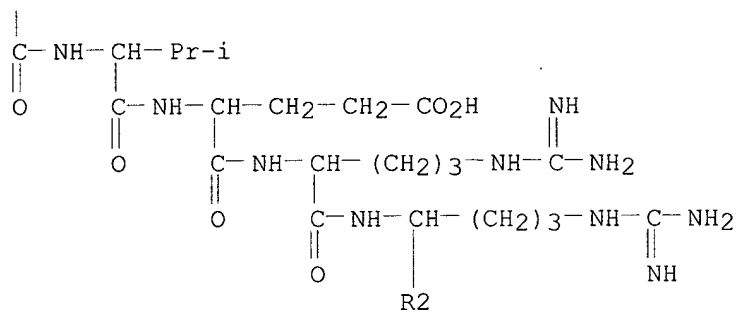


PAGE 2-A

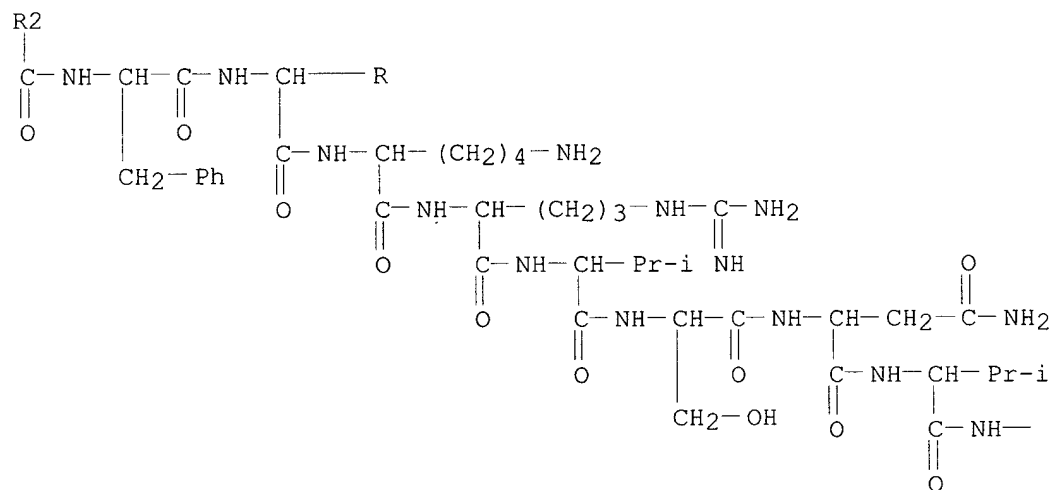




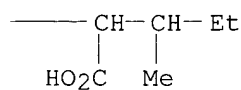
PAGE 2-B



PAGE 3-A



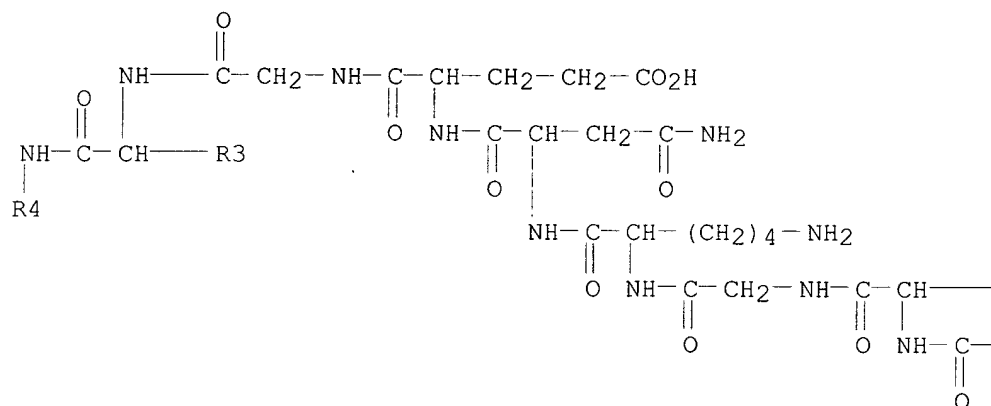
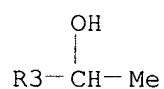
PAGE 3-B



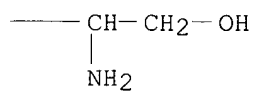
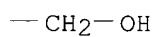




PAGE 4-A



PAGE 4-B



1 REFERENCES IN FILE CA (1962 TO DATE)  
1 REFERENCES IN FILE CAPLUS (1962 TO DATE)

REFERENCE 1: 127:277192

